A Functional Theory of Creative Reading*

Kenneth Moorman
Ashwin Ram

GIT–CC–94/01

January 1994

Abstract

Reading is an area of human cognition which has been studied for decades by psychologists, education researchers, and artificial intelligence researchers. Yet, there still does not exist a theory which accurately describes the complete process. We believe that these past attempts fell short due to an incomplete understanding of the overall task of reading; namely, the complete set of mental tasks a reasoner must perform to read and the mechanisms that carry out these tasks. We present a functional theory of the reading process and argue that it represents a coverage of the task. The theory combines experimental results from psychology, artificial intelligence, education, and linguistics, along with the insights we have gained from our own research. This greater understanding of the mental tasks necessary for reading will enable new natural language understanding systems to be more flexible and more capable than earlier ones. Furthermore, we argue that creativity is a necessary component of the reading process and must be considered in any theory or system attempting to describe it. We present a functional theory of creative reading and a novel knowledge organization scheme that supports the creativity mechanisms. The reading theory is currently being implemented in the ISAAC (Integrated Story Analysis And Creativity) system, a computer system which reads science fiction stories.

College of Computing
Georgia Institute of Technology
Atlanta, Georgia 30332–0280

---

*This paper is part of the Georgia Institute of Technology, College of Computing, Technical Report series.

Contents

1 Introduction

We read for a variety of reasons which differ from person to person and range from the straightforward and well-defined task of information gathering to the more amorphous goal of reading for pleasure. Reading seems to be tied to most of our day-to-day activities. Perhaps this is why so much interest has been given to the reading task in artificial intelligence, cognitive psychology, and education. Researchers have long been searching for a theory of reading. While many computer systems have been built to model this process and many theories have been proposed to explain it, few have come anywhere near succeeding. These earlier attempts failed for a variety of reasons which can all be traced to the fact that they did not make use of a complete reading theory. This resulted in these systems focusing on only portions of the problem. Systems which were primarily syntactic in nature, such as Woods' ATN systems (1986), did not understand text at a high level of comprehension. Systems which were successful at levels beyond the syntactic were limited to a small set of researcher generated texts and were unable to scale up to more realistic problem domains (e.g., BORIS, Dyer, 1983). Other systems (e.g., FRUMP, DeJong, 1982) tried for a wide coverage of texts but had to sacrifice high-level reasoning. Each attempt focused on only a portion of the reading process, all of which were too small to be successful. Comprehension is made up of numerous tasks which reinforce one another; as a result, focusing on too small an area results in a theory having to explain too much outside its scope, leading to overload and eventual failure.

We present a *functional theory* of reading designed to overcome the problems of earlier attempts. It is based on an analysis of the reading process, the knowledge required for successful reading, and an understanding of where earlier theories fell short. The theory describes the complete set of functions which a reader must perform during the comprehension process. Each function is modeled as a task. The various tasks maintain a close interaction, exchanging information as needed; this integrated approach lessens the burden on any one task. The theory relies on the huge amount of information available in a normal story; each task is dedicated to gathering and reasoning about a specific type of information. When one task discovers information, the entire system benefits. The tasks we have developed are sufficiently broad to describe the reading process as a whole, while allowing a researcher to see the impact and interestingness of each part, as well as the interactions between them.

We have also focused on an interesting type of reading we call *creative reading*. We are currently working in the realm of science fiction short stories, since these offer unique opportunities for this type of comprehension. 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 then they must understand the story itself. This is more difficult than understanding a story with a background that the reader is familiar with. We believe that all reading must contain some element of creative understanding; the reading process and creativity cannot be separated. Because of this, science fiction is methodologically a useful domain for reading research due to the fact that most stories will require at least some creative understanding to take place. We are implementing our theory in the ISAAC (Integrated Story Analysis And Creativity) system. ISAAC will be a comprehensive reading system; not only will it be able to read a variety of real texts, it will do so at a level of understanding never before achieved.

2 Reading

2.1 Formal definition of reading

Before continuing the discussion, we need to clarify exactly what we mean by the somewhat fuzzy term *reading*. In our research, reading is used to refer to the task of understanding a text of sentences and phrases which are related in some meaningful way. This text could be a novel, a newspaper article, a set of instructions, or the back of a cereal box. There are numerous other terms we could have chosen to use: *text comprehension*, *story understanding*, *interpretation*, *reading comprehension*, and *discourse understanding* are just a few well-known examples. Unfortunately, each of these terms has a set of connotations associated with it; furthermore, these connotations tend to vary across each discipline interested in studying the process of reading. To avoid possible confusion, we will use the less-used term *reading* to describe the overall cognitive task of natural language understanding.

Formally, we define:

> **Text**: A set of written sentences and phrases which are connected in some meaningful way to a common theme (or set of related themes). A text describes a set of scenes and episodes (scenarios) which are ordered temporally as well as causally. This ordering may be partial and implicit.

Our current definition for *text* limits its scope to the written word. We would stress, however, that many of the tasks which we will present will be useful in the understanding of *oral communication* as well. This is a hypothesis of ours which is beyond the scope of our current research; we merely wish to clarify the restriction of *written* in the definition of *text*.

> **Understanding**: The process by which a representation (usually tied to an external object) is transformed into an internal representation (which is usually some abstraction of the original) within a cognitive agent; while different from the original representation, the new one captures the salient features of the originally represented object. Successful understanding allows explanation and prediction to occur.

We can now define reading in terms of the above:

> **Reading**: The cognitive task of understanding a text.

For our particular purposes, the primary internal representation which is built as the result of understanding a text (the external representation discussed in the definition of understanding) is a set of three interlocking representations of the text—a representation of the events in the text, a representation of the structure of the text, and a representation detailing the reader's own processes that they went through to arrive at an understanding.

One final definition is required:

> **Creative reading**: The task of reading applied to texts containing concepts which are new to the reader.

Most reading will actually fall within the definition of creative reading; it is rare that all the elements of a text will be familiar to a reader. In order to read a text, the reader will need to perform *creative understanding* (Section 4.2.3) on the novel concepts. This is needed since the reader must understand the novel concepts and form a coherent interpretation of the text as a whole.

2.2 Example: *Men Are Different*

Consider the following story, *Men Are Different*, by Alan Bloch (1963), taken from a collection of short science fiction tales. The story is representative of several of the issues which a reading system must address.

> 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 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.

Men... illustrates several aspects of our earlier definition for a text. The sentences and phrases describe episodes and scenes which are connected in several ways; these include causal (e.g., the robot pushing the needle into the man's neck *results* in the man not moving anymore) and temporal (e.g., the events of the last two paragraphs occurred earlier in the life of the narrator than the beginning paragraph). Most importantly, the entire text is concerned with the common theme of misunderstanding the similarity between Men and Robots.

Reading *Men...* involves transferring the external representation of the text in English into a set of internal representations which capture the essential aspects of the story. This understanding includes story structure elements (e.g., the protagonist is a robot archaeologist and the setting is sometime in the future on a planet different from Earth), scenario elements (e.g., a robot archaeologist found a man and attempted to repair him when he was "broken"), and metareasoning elements (e.g., when did the reader understand that the narrator was not human or what the robot was going to do to the man). The reading and understanding of this text is an episode in itself—an episode of creative reading. The concepts of intelligent robots, the extinction of Mankind, and interplanetary colonies may all be novel to the reader, especially if the reader is unfamiliar with science fiction stories.

3 A functional theory of reading

Our theory of reading is a functional one. Any cognitive *process* may be explained by describing the *function* of each of its *tasks*, the *relationships* between them, the *mechanisms* which accomplish the tasks, and the *knowledge* which is needed. The resulting *functional theory* can then be used as a guide to the implementation of a cognitive model of the process in question. The theory itself does not explicitly tie the researcher to a particular method of implementation; instead, it describes the process at a higher level, the functional level. Using this methodology, it is possible to clearly see how a theory is broken into its *knowledge* components and *process* components. It is also easy to see exactly which processes make use of what knowledge in the system; this makes it easier to understand where the power in the system originates. In addition, it is possible to group related tasks together into a *supertask*; this helps identify and model the major cognitive behaviors of the system.

Using the above framework, one can see that what has traditionally been dubbed *story understanding* by the AI community is actually a set of interlocking supertasks which combine their abilities to produce a coherent understanding of a text. Each supertask represents a set of related tasks which together handle an aspect of the global reading process; it is only the sum of the supertasks which amounts to reading. Earlier story understanding systems can also be described as operating with a supertask/task breakdown; they simply concentrated on a specific one. The purpose of CYRUS (Kolodner, 1984), for example, was what we call the *memory management* supertask, ATNs were examples of our *sentence processing* supertask, and AQUA (Ram, 1991) concentrated on the *explanation and reasoning* supertask. These systems either ignored the other reading tasks or only implemented them at a level sufficient to demonstrate the performance of the ones being highlighted.

3.1 The reading task decomposition

We have identified six supertasks necessary for complete textual understanding. Each supertask is made up of numerous tasks which interact with each other within the confines of the associated supertask. A detailed discussion of each supertask follows:

- **sentence processing**: This supertask is responsible for low-level (mainly intrasentential) understanding, and includes tasks such as *pronoun reference*, *syntactic parsing*, and *lexical retrieval*. For example, in the first line of the last paragraph of *Men...*, a reader must retrieve meanings for the various words and determine who the "he" is referring to. Other tasks are *punctuation analysis*, which reasons about punctuation of phrases; and *tense analysis* which attempts to discover the text's tense. In many texts, low-level understanding issues will prove to be extremely helpful in gaining an overall understanding of the material. For example, the fact the "Men" is capitalized in the first sentence of *Men...* signals the reader that this is something special. Similarly, the italicized "really" in the second sentence is an emphasis which can aid the reader in understanding what is going on in the story.
- **story structure understanding**: This supertask handles the details of the text which relate to the story structure. Tasks include *character identification*, including protagonist and antagonist; *setting identification*, made up of time and location; *plot description*, which builds a coherent summary of the story's plot; and *genre identification*, which specifies in which category of text the story is. The inclusion of this supertask is largely the result of the way in which reading is taught in this country for the types of text which are most familiar in the Western tradition (Smith, 1986). While this area of the reading process may not be in-born or even learned at an early age, it does become automatic through our educational system. As such, it acts as a valuable aid in controlling and focusing the overall reading process. Realizing, for example, that *Men...* takes place in the future with a robotic protagonist acting as an archaeologist (in the truest sense of the word) greatly aids the reader in understanding the story.
- **scenario understanding**: The tasks making up the scenario understander are the *event parser,* which identifies various components such as agents, actions, states, objects, and locations; the *agent modeler*, which maintains descriptions of the agents, including their goals, knowledge, and beliefs; and the *action modeler*, which maintains descriptions of the acts with which the agents are involved. This supertask would be used in any experience in which the reasoner had to understand the actions of agents around it. As such, it is equally useful in day-to-day encounters with other reasoners as it is in the understanding of text. Generally, texts are intended to relate a scenario to the reader. *Men...* is not just a story, it tells the reader about a scenario which could potentially happen. A reader can understand the relationship between the robot and the Man in the story using some of the same processes which they would in an attempt to understand the relationship if it was occurring in the physical world.
- **explanation and reasoning**: This supertask performs high-level reasoning and learning. *Creative understanding* attempts to understand concepts which do not fit the reader's world view. The *unknown word definer* tries to use context clues, root words, and so forth to assign plausible meaning to words not in the lexicon. Consider a reader who is unfamiliar with the word *titanium*. Since this word occurs in *Men...*, the context will allow the reader to gain some comprehension of the unknown word. *Interest management* controls the reader's level of interest in the story. An avid science fiction fan, for instance, would be more interested in the example story than a fan of Westerns. *Belief management* reasons about the beliefs of the agents involved in the scenario. Did the robot actually believe that the Man was capable of being turned off? If so, how did it arrive at this erroneous conclusion? These are questions which will be handled by the belief management task as *Men...* is being read. The *explanation* task builds the inferences needed to connect the events of the story, enabling the reader to learn from the material. Finally, the *metareasoning* task reflects on the reader's own actions during the reading process; this information is also used for learning.
- **memory management**: This supertask handles general memory storage and retrieval, including spontaneous reminding. It is made up of *case building*, which constructs the various cases which result from a reading experience; *memory retrieval*, which returns information from memory which the system needs; and *memory storage*, which places new information and cases into memory. In addition to the conscious requests to store or retrieve elements in memory, the memory supertask should also handle the unconscious

storing of material and spontaneous remindings. During the reading of *Men...*, the reader will constantly be retrieving and storing concepts related to the material in the text. Some of these may be spontaneous; for example, the story may trigger memories of other stories which concern the extinction of humankind, or other stories written by Alan Bloch. Other memory processes will be more deliberate; the reader may decide, for example, to try to remember the last science fiction story they read about robots.

- **metacontrol**: Metacontrol integrates the other supertasks. It includes the tasks of *focus control*, which manages the depth of reading based on interest and understanding; *time management*, which allows the reader to make decisions based on time resources; and *suspension of disbelief*, which enables a reader to accept, at least temporarily, a text which violates their world view. This last function is particularly important in the case of reading a story containing unfamiliar concepts. A rational reader of the *Men...* example knows that the story cannot be true. Mankind still exists, space travel doesn't, and today's robots are merely mechanical devices used on assembly lines. In order to read, understand, and enjoy the story, however, the reader must be willing to accept these unfamiliar ideas for the duration of the reading experience. Other aspects of metacontrol are important in all reading episodes, whether creative or not. Readers are cognitive agents involved in numerous ongoing tasks. They must decide how much effort to expend on each task as a whole, as well as on each part of an individual task. As part of this resource management, the reader must also decide in how much depth they wish to read the given material. With *Men...*, for example, an average reader will not read each and every word; instead, some sections will be read closely, others will be skimmed, and some may be skipped entirely.

The high-level organization of the six supertasks is illustrated in Figure 1. Natural language text enters the metacontrol supertask, which directs the other supertasks. The three "primary" supertasks (metacontrol, scenario understanding, and story structure understanding) produce structures which capture the understanding of the story (what the reader did to understand the story, what events occurred in the story, and what the structure of the story was). The three "support" supertasks assist the primary ones in the understanding process.

The supertasks were designed based on functional analyses of introspective reasoning about our own reading processes, backed up by a significant amount of prior research in the areas of psycholinguistics (e.g., Holbrook et al., 1992; van Dijk & Kintsch, 1983), reading comprehension (e.g., Black & Seifert, 1981; Graesser et al., 1991), story understanding (e.g., Birnbaum, 1986; Ram, 1991; Rumelhart, 1977), memory (e.g., Kolodner, 1984; Schank, 1982), and metacognition (e.g., Gavelik & Raphael, 1985; Wellman, 1985; Weinert, 1987; Schneider, 1985). While metacognition is important, many psychological results have shown that humans do not have total access to internal processes (Nisbett & Wilson, 1977). In our model, metacontrol and metareasoning only have access to supertasks, not to individual tasks. In addition to providing a cognitive model of reading, our task decomposition allows a straightforward analysis of past systems; most previous systems concentrated on one particular supertask in our theory, which led to an inability to model the complete reading process.

3.2 The model-driven aspects of the theory

In addition to the modular breakdown of reading tasks, our theory makes use of *models* to drive the comprehension process. A model contains knowledge concerning what a supertask should expect from different types of texts; specific types of models exist for each supertask in our theory. These models allow flexible behavior and the ability to read from a wide range of texts. Humans learn what to expect from different types of stories; this knowledge aids them in understanding later stories. To illustrate, if a reader has read a number of books by Alan Bloch, they will develop a specific *story structure model* regarding what to expect from the next Bloch story, or even the next story claiming to be written in the Bloch style. This is in addition to any models which may exist concerning how to read science fiction or fiction in general. Examples of story structure models include genre specific models, as well as models describing narrative types (such as first-person, second-person, and third-person narratives). A similar idea has been proposed by Carpenter and Alterman (1993); in their model, case-based reasoning is used to drive the reading process which "knows" how various text genres should be read (e.g., when confronted with a set of complex instructions, skim the first few pages looking for an enumerated list as these are particularly important to instruction reading).

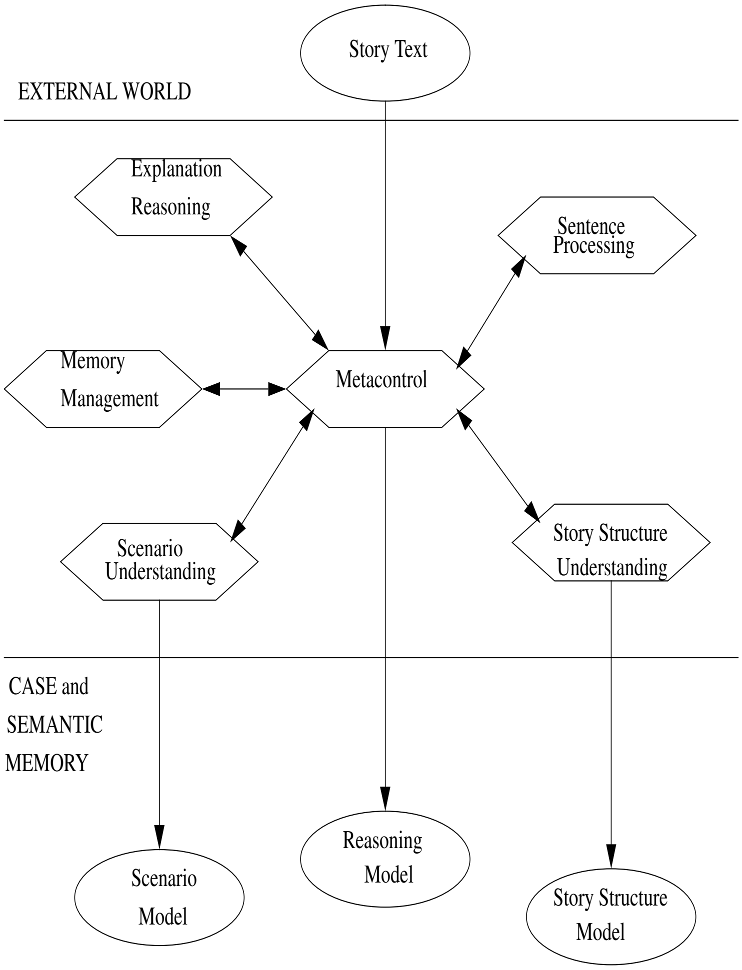

Figure 1: Functional model of reading

Scenario models explain different styles of agent interaction. This includes various combinations of "social states" in communication; for example, the type of interactions one would expect to see between two close male friends would be very different from the type of interaction exhibited between a boss and an employee in a formal business setting. Several of these social relations can be seen in *Men Are Different*: the protagonist is a peer among the robot archaeologists, it was a student of the historians discussed in the story, and it was a researcher with respect to the experimental subject (the man). Augmenting the social and interaction models are models of how agents react in various emotional settings (e.g., if a person receives poor service in a clothing store he regularly visits). Finally, "local" scenario models exist which describe the more-or-less default results from an action in a standard setting. At the lowest level, these models are simply the expected results for a given action and the expected preconditions which may have triggered the action; as such, they allow low-level inferencing to be performed (Schank et al., 1975). At their highest level, the local scenario models resemble scripts (Schank & Abelson, 1977); they package together a set of expected preconditions and results which exist in various stereotypical events (such as going to a fast-food restaurant; or, as in the *Men...* example, the process of repairing a malfunctioning "device" by using field spares).

Next, *sentence models* give the capability of dealing with different low level structures. Questions, declarative sentences, quotations, broken or interrupted speech, and sentence fragments are examples of these types of models. The reader's high-level reasoning is guided by the various *reasoning models*. These control, for instance, how much explanation a reasoner may choose to do in a given scenario; if a person's car breaks down on the highway, do they immediately try to find help, reason and examine the immediate possible causes, or attempt to reason the problem through to completion? *Metacontrol models* allow the reader to tailor their reading patterns to a given situation. Is the material being read for pleasure, for study, or to kill time? Lastly, *memory models* permit different types of memory access to occur based on the text involved. Some situations may trigger deeper remindings than others. Based on the type of material being read, a reasoner may request the construction of a careful memory of the material (e.g., studying for a midterm test) or request a more shallow memory be built (e.g., reading a cereal box for the list of ingredients). These requests would all be handled by various memory models.

To see the power of these models, consider a reader who realizes that they are dealing with a first-person science fiction story. They will read it differently than if they think they are dealing with a third-person newspaper article. The focus in a newspaper article is to build coherent descriptions of events. The reader can expect the language to be more concise than in a narrative. Also, not as many inferences need to be made since the purpose of a news article is to inform in a clear manner. With the science fiction story, on the other hand, the reader must take more of an active role. The story will probably contain concepts which are unfamiliar to the reader. A reader must decide how much effort they will expend in the attempt to understand these new concepts. Accurate character models and models of character interaction must be built and maintained if the story is to be appreciated. The first-person aspect of a story like this lessens the burden of discovering who the protagonist is and enables the reader to concentrate on other story aspects.

3.3 The metacognitive aspects of the theory

As mentioned earlier, our theory includes mechanisms for metacontrol and metareasoning. This means that a reader is aware of not only the text being processed, but also of some of the internal workings of their own cognition during reading. Metacontrol is responsible for the division of labor within the system. Based on the current interests of the reader, the available cognitive resources, the purpose of the reading episode, and the difficulty of the text, the metacontrol supertask decides what other supertasks need to be executed and on which portions of the text. Humans obviously do not read each word of a text to the greatest depth of understanding possible. No reader, for example, would have the time, memory, and other resources to read every single word of a story (even a short story) and consider the possible extrapolations and extensions resulting from each. The reader's current environment, knowledge, goals and tasks (Ram & Hunter, 1992), and cognitive resources (e.g., Just & Carpenter, 1992) interact to control the level of understanding which is attempted. Therefore, a comprehensive theory of the reading process must also explain when and why in-depth reading occurs, as well as when and why shallow reading, such as skimming or skipping text, occurs.

The metareasoning portion of the theory enables the reader to reflect on their actions during the process. Remember the last mystery story you read? When did you realize who the killer was? What clues tipped you off? Did your opinion of the person change as a result; if so, why did it change? These are all the kinds of questions which the metareasoning task will enable the reader to answer. Metareasoning results from keeping explicit representations of the reasoning processes occurring during the cognitive activity. These *reasoning traces* can then be reasoned about, much like any other piece of knowledge within a cognitive agent (Ram & Cox, 1993). This functionality enables a reader to reflect on their own reasoning, thus enabling them to learn more from the experience. For example, if the reader discovers that they were "led down the garden path" at some point in the story, the metaanalysis will enable the reader to learn this and therefore (potentially) avoid a similar problem in the future.

It is also possible to view metareasoning as an ends as well as a means as in the previous examples. For example, the enjoyment in reading most mystery novels is not derived from the conclusion; otherwise, the author could simply inform the reader upfront as to "who-done-it." Rather, the enjoyment comes from the twists and turns that the reader is forced to endure during the process. The final discovery of the true criminal results in a sense of supreme satisfaction, as one can look back over their reasoning patterns and see exactly how the conclusion was achieved.

4 Creative understanding

It could be argued that no system can accurately implement all the aspects of our large theory and that the only way to build reasonable systems is by restricting the coverage. We believe, however, that there are better ways of restricting the problem than by arbitrarily focusing on a subset of reading tasks, as many earlier systems did. Many interesting issues can be explored by concentrating on a particular genre of stories and then developing a computational model of the complete task decomposition to support the reading and understanding of these stories. By concentrating on the portions of the theory which most directly support the creative understanding features, we make the research problem more manageable; by maintaining an awareness and an understanding of the entire reading process, we avoid the pitfalls of earlier systems. Our theory, while it describes how people read in general, is primarily concerned with how people read *creatively*.

4.1 Working definition of creativity

People have preconceived notions of exactly what creativity is and how to judge it in other people. Generally, most people see creativity as something unexpected which is brilliant in some unexplained way. Or, creativity may be seen in the artistic works of an artist; a result which possesses some sort of beauty or relevance never existing before. Unfortunately for AI research, these feelings and instincts concerning creativity are hard to operationalize. A definition of creativity is needed which precisely defines its aspects. We propose the following candidate:

> **Creativity** : A directed, internal process (Section 4.1.1) of a cognitive agent which results in an artifact (Section 4.1.2) which is both novel (Section 4.1.3) and useful (Section 4.1.4).

The definition is not intended to capture every nuance of what makes an act creative; it is also possible that some of the processes which fit this definition may not be considered creative by society. Still, it serves its intended purpose of providing a common definition of the creativity phenomenon, allowing us to examine closely exactly what our processes need to do to fall within this definition's boundaries.

4.1.1 Process

In order for a process to be creative, it must be both *internal* to the reasoner involved and it must be *directed*. The internal restriction ensures that the reasoner is not simply repeating a piece of knowledge they just received from another source. For example, if an outside person tells a reasoner her phone number, the reasoner then possesses a bit of information which is novel to them and which is certainly useful (if the reasoner in question intends to contact the person). No one, however, would argue that the reasoner has performed something creative. They did not internally produce the phone number in their mind, they merely acquired it. For a result to be viewed as creative, it must have been internally generated by the cognitive processes of the reasoner.

The other restriction, that of directedness, ensures that the agent in question did not simply randomly generate solutions until a creative one arose. While randomness is often seen as a necessary part of creativity (Boden, 1991), it is not desirable for this randomness to be totally unconstrained. The old AI adage concerning an infinite number of monkeys typing on an infinite number of typewriters for an infinite span of time is appropriate here: while they will certainly generate an infinite number of creative artifacts, the process by which this generation is accomplished is not directed. Directedness is an important aspect for determining the creativeness of a given process.

The presented working definition treats creativity as a process. While it is true that a creative process can only be judged by the artifacts it produces, as AI researchers we are interested in identifying the processes responsible for this generation. It is our belief that a set of mechanisms exist which interact to produce the phenomenon popularly referred to as "creativity". Some of these mechanisms will exist in so-called mundane activities while others will appear primarily (if not exclusively) in the activities identified as creative. It is therefore important to attempt to recognize these mechanisms of potential creativity, to identify the knowledge which each must have access to, and to see the types of interactions which are possible.

4.1.2 Artifact

In this formalism, *artifact* is defined to be the result of a mental process by a reasoner. This result may be an object, which can either be physical or mental. Alternatively, the artifact may be another process. This is the case when a reasoner uses a mental process to develop a second process.

An artifact is described by a set of *attributes* which define important characteristics of the artifact. One of these attributes is the *function* (or set of functions) which an artifact possesses. This may not represent the complete functionality of the given artifact; most artifacts can potentially be used for hundreds of wide varying functions. Instead, the function attribute represents the traditional or best-known uses of a given artifact. The remaining attributes can be divided into primary and secondary ones. *Primary attributes* are those which contribute to the explanation for why the artifact can perform the function which it has; *secondary attributes* are the remaining ones. Each known function of an artifact configures the attributes into potentially different primary and secondary sets.

4.1.3 Novel

Novel is a more difficult term to define. Idealistically, an *N-artifact* (novel artifact) exists whenever an artifact exists which was previously unknown to the reasoner. But, there are degrees of this novelty. For example, if the reasoner only knows about *wooden pencils which are eight inches long* and them reads about a *wooden pencil which is ten inches long*, it is technically an N-artifact. However, this seems somehow less novel than if the reasoner reads a description of a pencil which is a *mechanical pencil which is eight inches long*. In the first case, the pencil is still a wooden pencil, its length has changed. This probably does not affect its ability to be a pencil. The second case, however, represents a new type of pencil, one which may potentially be a better pencil than the original example. Yet, both are new to the reasoner. It should be clear that a more detailed view of novelty needs to be considered.

Novelty of an artifact is defined with respect to a particular point of view (of some reasoner or set of reasoners) and with respect to the functionality of the artifact in question. There are three broad ways in which an artifact (M) may be novel with respect to a given function or goal (G):

- Absolute Novelty (A-Novel): M is defined to be A-Novel iff M is unknown to the reasoner whose point of view is being considered.

- Evolutionary Novelty (E-Novel): M is defined to be E-Novel iff M is A-Novel and M accomplishes G in a better way than other examples of artifacts which accomplish G. This is generally the result of altering one of the primary attributes of M.

- Revolutionary Novelty (R-Novel): M is defined to be R-Novel iff M is A-Novel and M accomplished G in a new way than other examples of artifacts which accomplish G. This form of novelty requires more extensive changes than to simply alter the primary attributes of M. Secondary attributes may be altered in ways which cause them to now participate in the function of M, attributes may be added to either set, attributes may be removed, and so on. The function of the artifact remains the same; the underlying explanation of how it accomplishes this function is what has altered.

To illustrate these differences, consider a black longsword. Without question, a red longsword will perform the same function as the black longsword. So, the red sword, in this scenario, would be described as A-Novel. Within a range of possibilities, a short sword, a bastard sword, and a two-handed sword are all capable of performing the same function as the original longsword. Therefore, these are artifacts which are E-novel. Finally, the light saber from the Star Wars series is also capable of performing the same function as the original longsword. In this case, the light saber is an example of R-novelty.

With regard to point of view, one can see that there are two dimensions to consider, an internal perspective and an external perspective. If an idea or act is neither novel from the perspective of the reasoner or from the perspective of the public, it is a totally mundane act. If the act is creative to the individual but not to society in general, it is a self-creative act, also known as *psychological-creative* or *P-creative* (Boden, 1991). An example of this would a child discovering the number-theoretic relationship that the digits of a number divisible by nine must sum to a number divisible by nine. To a child, this is a fantastic discovery; to a mathematician it is a well-known rule. The reverse of this is when the reasoner sees an act as mundane but others view it as creative. This may come about if the reasoner has some knowledge that the general public lacks. Finally, a *historical-creative* or *H-creative* (Boden, 1991) act results if the reasoner finds an act creative as does the rest of society.

In both the internal and external perspectives on novelty, it is important that the reasoner whose point of view is being considered is aware that the artifact is indeed novel . A system, for example, that possessed no memory of its own actions could conceivably be considered to be producing nothing except internally novel artifacts; after all, each new one is novel from the reasoner's perspective. In our view, this memory-less reasoner would be incapable of producing internally novel artifacts since it would be incapable of realizing their novelty.

4.1.4 Useful

Useful is perhaps a bit easier to define for an artifact. A reasoner attempting to perform creatively will have a certain task to accomplish and a certain goal to achieve. In order for a process to be considered creative, it must produce a solution which fulfills the goal and accomplishes the task possessed by the reasoner. This restriction does not say that the solution must be optimal for the process to be considered creative, just that it must be useful. The usefulness aspect is straightforward to define in our chosen domain of reading: understanding novel concepts within a text is intrinsically useful to a reader. For example, understanding the functioning and purpose of the light saber is useful to the reader of the Star Wars series because the understanding helps the reader enjoy the story more.

4.2 Functional definition of creativity

The provided definition of creativity is descriptive of the process. Given a particular process in a certain situation, it would now be possible to judge whether that process is mundane or if it exhibits some level of creativity. However, details concerning how this view of creativity is operationalized still need to be presented. A more formal explanation of creativity is needed in order to understand the implementational details associated with it. In particular, it is necessary to identify the mechanisms and knowledge which is needed to accomplish processes which can be viewed as creative.

4.2.1 Formalization of problem solving

In the abstract, creativity can be seen as an extension to what is traditionally called *problem solving* by the cognitive science community. Historically, problem solving begins with the reasoner in an *initial state* (or *start state*). A reasoner knows of various operations that they may perform which will move them through a *search space*. This search stops when the *goal state* is achieved. The output from the process is a *solution path* which takes the reasoner from the initial state to the goal state (Newell & Simon, 1972). Later research showed that the idea of *constraints* was an important one (Sacerdoti, 1974; Sussman, 1973). These are conditions which cannot be violated in the final solution to a given problem. Finally, there are times when the reasoner may also possess a solution to the problem in question. In this case, the role of problem solving can be to discover a better solution than the one known. To accomplish this, the reasoner needs to possess a critique as to why the current solution is not a viable one (Hammond, 1989). The final algorithmic formulation of the problem solving process is shown in Figure 2.

With this formalization of problem solving in place, we can now examine the phenomenon of creativity in more detail. Roughly, creativity can be divided into two large areas, invention and understanding. Each plays an important role in intelligence. Although we are primarily interested in creative understanding, there is much in common between the two areas. So, it is beneficial to examine both aspects of the total creativity process.

FUNCTION **Problem-Solver**

- INPUT:
    - Initial state (I)
    - Goal state desired (G)
    - Set of constraints (C) (optional)
    - Current solution (S) (optional)
    - Critique for why S is not a good enough solution to stop with (K) (optional)
- OUTPUT:
    - Solution path (S') which achieves G given I and does not violate C

Figure 2: Formal algorithmic description of problem solving

FUNCTION **Understander**

- INPUT:
    1. Goal state desired (G) AND/OR
    2. Current solution (S)
- OUTPUT:
    1. A possible solution (S) OR
    2. The goal state initially desired (G) OR
    3. Critique of S as a solution to G (K)

Figure 3: Formal algorithmic description of understanding

4.2.2 Creative invention

Creative invention (also called *creative design*) represents the generative side of creativity (for a recent discussion of creative design, see Kolodner & Wills, 1993). When a reasoner is faced with the task of designing a novel artifact (as defined in Section 4.1.3), they accomplish it through creative invention. When Motorola's engineers developed the MicroTAC, they were engaged in creative design. Not only did the designers have to create a portable, pocket-sized cellular phone, they had to deal with the fact that people often drop things that small. Before the design was complete, the team had helped to push the plastics industry far enough ahead to actually provide the material that the MicroTAC needed. As a result of this and other innovations in the phone's design, it was recently highlighted in an issue of *BusinessWeek* (1993).

4.2.3 Creative understanding

Before discussing *creative understanding*, it is necessary to realize that a formalism similar to the problem solving description in Figure 2 can be applied to understanding in general. The understanding formalism is shown in Figure 3, which presents all of the optional inputs and matching outputs. In general, understanding can refer to one of three things. Given S, understanding should tell the reasoner what possible G's are. If a reader sees a robotic story character try to turn off a man and then open him up, they may understand the scenario by reasoning that the robot had the goal of repairing the man and probably thought he was malfunctioning. This process of reasoning "backwards" from the solution to the beginning is also called *abduction*. Second, given G, understanding should tell the reasoner what possible S's are. A reader who learns that Mankind has become extinct and that the remaining robots are curious as to the final fate of their creators may understand this by reasoning about upcoming story actions. Better understanding leads to better *prediction*. Finally, given both S and G, understanding should provide an explanation of *why* S is a good solution to G. In the specific example of *Men...*, the reader must understand why the robot felt that performing field repairs on the man was a good solution to his temperature discomfort. This aspect of understanding is sometimes referred to as *explanation*.

Mundane understanding is more a recognition task than anything else. If a reasoner knows explanations relating goals and solutions, and the reasoner is then presented with both a goal and a solution to explain, the problem is simply recognizing the configuration and retrieving the proper items from memory. Similarly, a reasoner may know what goals and solutions are logically linked; when presented with one, it is fairly straightforward to retrieve the corresponding other. On the other hand, *creative understanding* is the process by which new concepts are understood by a reasoner. This is a more difficult task than mundane understanding in the sense that there are no in-memory items which directly answer the understanding questions. As a result, the reasoner must resort to creative techniques in order to attempt to understand the novel concepts.

For the task of reading, consider the fact that people are able to comprehend stories which have no basis in fact, as anyone who has ever read a *Star Trek* novel can verify. People are also able to transfer concepts which exist in a story world into real world ideas. This idea of the creative level of understanding is directly related to the idea of an *active reader*. An active reader will tend to engage the text more and have more opportunities for creative understanding. In fact, many researchers refer to active reading as creative reading (e.g., Robeck & Wilson, 1974; Gruber & Davis, 1988). Our research has led us to believe that creative understanding is a key component of the overall reading process and that any theory of reading must explain it satisfactorily. Thus, creativity is an essential and integrated aspect of our complete reading theory.

5 Past approaches to creativity

While there is little research expressly concerned with creative understanding, there have been many attempts to explain exactly how the phenomenon of creativity occurs in the human species; these researchers have generally focused on creative invention when they discussed creativity theories. A number of theories, unfortunately, chose to concentrate on describing the process without really explaining how it occurs. Without such an explanation, it is difficult to develop a model of the cognitive processes which act to produce creativity. It is still beneficial, however, to examine some of these descriptive formalizations of creativity in order to gain insight into what a good description of the process might be.

Perhaps the best known formalization of the creative process comes from the work of Wallas (1926). Wallas viewed creativity as being made up of four steps: preparation (the reasoner gains a rich knowledge of the problem), incubation (the reasoner's subconscious is working on the problem), illumination (a potential answer is consciously revealed in a flash of inspiration), and verification (where the reasoner examines, corrects, and refines the solution). This points out that creativity takes a great deal of knowledge and preparation and that the unconscious plays a large role in the process. Beyond that, it does not reveal much regarding how the process would be carried out or exactly what makes up each step. As a result, it provides a good description of the creative act but does not help us to understand it to the point where it may be modeled.

Weisberg (1986) produced an in-depth criticism of Wallas' work on creativity. His description of the process does not rely on the near-mystical aspect of illumination found in earlier approaches; instead, creativity is an extension of normal problem solving.

- Creativity is an extension of problem solving using the same processes in more flexible ways.
- There is no single "creativity" module in reasoning; instead, creativity arises from a complex and interconnected set of processes, including problem solving, recall, adaptation, and so on.
- People do not either possess creativity or not possess it; rather, creativity exists everywhere in an incremental continuum from the mundane to the more and more creative.

While more open-ended than the Wallas account, Weisberg still shys away from attempting any more defined model of the process. Implementational details are, as textbooks often say, left to the reader.

One place to look for explanations of how to perform creatively is in works dedicated to improving the reader's own creative abilities. While these are not generally intended to be valid psychological models of creative processes, they can give insight into what sorts of processes need to be considered. Two such techniques designed for improving people's creative abilities are *attribute listing* and *morphological synthesis*.

Attribute listing, Crawford's work (cited in Finke et al., 1992), begins with a reasoner considering an object which they think can be the springboard to a new, creative artifact. Each of the object's primary attributes are examined, with the reasoner thinking about ways in which each could be manipulated. Finally, the reasoner filters out the manipulations which lead nowhere; those that remain are considered to be creative.

Morphological synthesis, Allen's technique (cited in Finke et al., 1992), begins in the same fashion as attribute listing. The reasoner considers both attributes and commonly accepted values for each. Then, new combinations of attributes and values are examined. The reasoner looks for unknown combinations which may prove to be creative. While this is similar to attribute listing, in morphological synthesis, the reasoner does not concentrate on changing attributes one-at-a-time; instead, interesting combinations are concentrated on. This leads to faster generation of potentially creative artifacts.

Consider a piece of luggage, an everyday mundane sort of object. To perform attribute listing on it, a reasoner would list all the *primary attributes* of the object and their prototypical values; in essence, the reasoner would bring to mind an "average" piece of luggage. It is made of leather, about two foot by two foot in size, has a soft body construction, one large compartment for storage, and is carried by means of a handle. After these major attributes are listed, the reasoner starts to consider altering them, one at a time. Some possible changes are: *cloth* may be substituted for *leather*, the dimensions may become larger or smaller, the the number of compartments may change, or *wheels* may be used instead of a handle. The reasoner may have to "clean up" the new product in some situations; for example, if the number of compartments become two, the reasoner must decide how the two compartments will be integrated. Morphological synthesis would go about the process in much the same manner, with interesting combinations of attributes considered rather than single ones. The idea is the same as in attribute listing, the focus is just a little larger. So, the reasoner may decide to consider a large piece of luggage with lots of compartments versus a small piece of luggage with only one. As more attributes are added to the set under consideration, the number of possibilities grow.

The obvious drawback to these approaches is that a reasoner will never develop an R-novel object, just relatively close variations of the original concept. As discussed in earlier, this is due to the limitation of only considering primary attributes. As a result, E-novel artifacts are the only ones which can be created. The main advantage of morphological synthesis is the capability of attribute combinations; this decreases the time needed to produce potentially creative artifacts.

Notice, for instance, that neither of the two allow the introduction of new attributes. Yet, one could imagine a number of additions to a piece of luggage which would be useful. Many travelers, for example, forget to bring an alarm clock with them. One could conceivably add a small alarm/radio to a piece of luggage; since it is unlikely that a traveler would forget their luggage, an alarm clock would always be "remembered." Or, a mirror could be added to the inside of one of the compartments. Another possible addition is the idea of dedicated compartments of the suitcase being used for specific items.

6 Computational model of creative understanding

Creative understanding, as we have defined it, has a well-defined input and output. The research problem to be addressed is to identify the set of mechanisms which accomplish the function as a whole. Thus, creative understanding will be broken down in the same manner that the entire reading task was in Section 3—a functional decomposition. In order to achieve this, both the relevant processes and the required knowledge needs to be identified. Prior to the discussion of the mechanisms and knowledge required for creative understanding, there are certain requirements for creativity, in general, which need to be explained.

6.1 The willing suspension of disbelief

The first requirement for creative understanding is a willingness by the reader to suspend their skepticism of the material being presented. This *willing suspension of disbelief* (first described by Samuel Taylor Coleridge as reported in Corrigan, 1979) occurs in practically all reading, although it is more prevalent in more fanciful texts. Consider the ambiguous title of a Larry Niven story, *Flight of the Horse* (1973). This phrase could refer to a horse on an airplane flight, to a fleeing horse, or to a flying horse. If a story understanding system relied on a belief in the validity of world knowledge (as most previous reading systems did), it would disambiguate the phrase to rule out the latter meaning since it "knows" horses cannot fly. This is wrong if the story was about a flying horse, which is perfectly reasonable in a science fiction story.[1] The more familiar a reasoner is with the concepts of a story (or similar concepts), the less suspension of disbelief will be required. This is one reason why a reader finds it easier to read texts in genres in which they are most familiar; they are already familiar with a large number of concepts which makes new concept understanding easier than it would be for a novice in that genre.

6.2 Relation of invention to understanding

Although we are interested primarily in creative understanding, there is an interesting approach to it which depends on creative invention. One way for a reasoner to creatively understand a concept is to discover how they would have creatively invented the concept in question, a sort of reverse engineering of the creative act. The reasoner begins with the novel concept and manipulates it until it resembles a concept they are more familiar with. This leaves them with a solution path which, if reversed, would creatively invent the concept it started with. It is not necessary that the concept being understood was actually created in the way that the reasoner comes up with. An example of this is the unicorn (example results from a dialogue with Anthony Francis (1993)). The modern view of the unicorn is a white, horse-like creature with a single horn. Various attributes are given to the beast, including nobility and purity. A reasoner who is ignorant of unicorns may attempt to understand a reference to one by appealing to their knowledge of existing creatures, beginning with the closest analogue, the horse. Knowing that horses are capable of defending themselves through the use of hooves and teeth, the reasoner may then combine the horse with an animal possessing a horn (or horns) used for defense. The general shape of the unicorn is now in place. The reasoner may further make use of their knowledge of what the color white generally symbolizes to make the inference that unicorns are pure and good. This is reinforced by the reasoner's impressions of horses as generally "good" creatures.[2]

This is a perfectly acceptable way of creatively understanding the concept of unicorn by imagining what it would take to creatively invent such a creature. However, this is not the way in which the concept of unicorns was initially conceived. The unicorn was a mythological animal and was initially seen as the blending of several existing animals. It had the body of a horse, the tail of a boar, the feet of an elephant, the head of a deer, and a two cubit long black horn in the middle of its forehead (Bulfinch, 1959). The description was developed over a period of time through the oral tradition of the Greeks. It was based on glimpses of animals, a willingness to believe in the supernatural, and the conceptual changes which occurred as one traveler relayed his story to another, who told another, and so on (much like the children's game of Telephone or Gossip). Over time, the

---

[1] Purists may object to a definition of science fiction which allows pegasi to be included; many would prefer a stronger division between science fiction and what might be more traditionally called fantasy. Science fiction generally takes the current world, focuses on one aspect, and then projects a "what-if" scenario considering the world if that one aspect was changed in some significant way. Fantasy generally creates an entirely new world, usually with only a minor relationship to the current world. Fantasy, rather than utilizing technology, generally makes use of magic. We have are currently choosing to focus on those stories which resemble the current world with some changes. As a result, the majority of our focus will fall in the realm of science fiction, although that is not necessary. A story involving a pegasus, for example, is not any intrinsically harder to understand than a story involving an alien creature. As far as the magic aspect goes, this is also a non-issue; as Arthur C. Clarke noted, any sufficiently advanced science is indistinguishable from magic (1962). Our working definition of science fiction for the current stage of the project are those stories which stretch the "real" world in some significant way while still maintaining a good deal of similarity with it. As a result, this will include a great deal of science fiction and fantasy both.

[2] As a counterpoint, a reasoner presented with the unicorn-like creatures in Bailey's science fiction novel, *Brothers of the Dragon* (1993), would be lead to a very different concept of unicorn. These unicorns are not the merging of a horse with some horned animal, but rather the merging of a lizard-like or snake-like creature with a horned animal. They are armored with edged-scales and jet-black in color. Thus, rather than theorizing that these unicorns are noble and good, the reasoner would be more likely to see them as evil and base.

description focused more and more on the horselike qualities and the other definition faded. Also over time, the concepts of purity and nobility were added. Bullfinch (1959) notes that this was probably an attempt to explain how the beast might be captured; since it embodied purity, an innocent could conceivably lure the creature into a trap. A reasoner does not need to understand the historical development of the unicorn mythos in order to understand a unicorn; they do need creativity.

6.3 Uses of a creatively understood story

After a story has been creatively comprehended, there are two ways it can directly aid the creative process. First, it can act as a source of ideas. For example, reading about a certain device may prompt the reader to develop a similar one in the real world. While the mind of the actual design team is unknown to the authors, the Motorola MicroTAC cellular telephone bears an uncanny resemblance to the communicators used on the original *Star Trek* series. The lapel-style communicators seen on *Star Trek: The Next Generation* strongly resemble the concept of SmartBadges as proposed and implemented at the XeroxPARC site (Weiser, 1991). While these similarities may be simple coincidence, the June 7 issue of *BusinessWeek* points to a stronger real-world example of this phenomenon. The designer of Hewlett-Packard's new portable DeskJet (an inkjet printer) attributes a major design inspiration to Gort, the robot servant in the 1951 classic science fiction movie, *The Day the Earth Stood Still* (BusinessWeek, 1993). Current research in this paradigm includes the work of Goodman, Waterman, and Alterman (1991). Their SPATR system can reason about the use of novel and potentially confusing devices (such as the AirPhone) by appealing to its existing information regarding the use and functioning of known devices (such as a payphone or an automatic teller machine). The other way in which reading assists creativity is similar to Carbonell's derivational analogy (Carbonell, 1990). If a reader observes a character problem solving creatively, they may be able to replay the solution process on a real problem in order to develop a creative solution.

6.4 The "creative" in creative understanding

We can now turn to the issue of exactly *how* creative understanding is to take place. Descriptively, creativity in general and creative understanding in particular are bound by the following constraints:

- Creativity arises out of the same sorts of processes as does mundane reasoning.
- It is possible to identify some processes as being more open to creativity.
- People who exhibit creativity generally have a great deal of knowledge about the domain in question.
- Creativity requires a willingness to take a cognitive risk, to suspend disbelief while exploring new concepts.
- Creativity cannot be objectively evaluated; the situation often determines whether or not something is tagged as creative.

For the explanatory side of our creative understanding theory, we are working under the hypothesis that it is possible to identify a core set of cognitive mechanisms which contribute the most to creative processes. We begin with our algorithmic description of the understanding process, introduced in Section 4.2.3. Since the reasoner is dealing with new concepts not in their memory, there are three basic ways which they may attempt to understand them. First, analogy may be employed. It may be the case that a new concept is sufficiently like a known concept that a process of analogical reasoning can be employed. The reasoner's understanding of the known concept is used to understand the new concept. The second approach is to reason from general knowledge which the reasoner possesses. This is the application of abstract knowledge to the problem of understanding the new concept. Lastly, the reasoner may attempt to creatively reverse engineer the new concept to something they are familiar with.

```
FUNCTION Creative Understanding Process (CUP)
    • INPUT:
        – Solution (S)
    • OUTPUT (desired):
        – Goal state desired (G)
        – Constraint set (K)
    • PROCESS:
        – REPEAT
            1. Attempt to perform normal memory retrieval (6.5.1)
            2. If an item is returned by (1) and it is not an exact match ⇒ attempt analogy (6.5.2)
            3. If (2) fails ⇒ attempt baseless analogy (6.5.3)
            4. If (3) fails ⇒ reformulate the problem (6.5.4)
        – UNTIL (successful OR iterations exceed a threshold)
```

Figure 4: The Creative Understanding Process

6.5 The creative understanding process

The three basic ways to achieve creative reasoning can be combined into one algorithm, the *creative understanding process* (CUP). The cycle of creative understanding is shown in Figure 4 (while the description indicates S as an input and G as the desired output, the same approach is used for the other examples of understanding discussed previously). Mundane understanding exists if the reasoner only considers steps 1 and 2; that is, if the reasoner only attempts memory retrieval and analogy. Each cycle through the complete function, on the other hand, increases the potential for creative understanding. At some point the reasoner will be so far removed from the original concept attempted to be explained, that further iterations will be useless. As will be discussed in Section 7, there is a point (dependent on the context and background knowledge of the reasoner) where creativity degenerates into bizarreness. Still, there is no theoretical limit to the "amount" of creativity which may be generated; it is potentially unbounded. As discussed earlier, the reasoner makes a decision to suspend disbelief of potentially creative concepts in order to explore the possibilities they represent.

6.5.1 Imaginative memory

The first step of the CUP algorithm involves a memory retrieval attempt. This is intended to be a more-or-less normal retrieval from memory, although it can contain elements of imaginative memory response. *Imaginative memory* was a term used by Turner (1992) to describe the creative mechanism within his MINSTREL system. This technique places the responsibility of the alterations to known objects in the control of the memory system. Suppose a reasoner is attempting to remember a scenario in their past dealing with knights killing dragons. Their memory contains no such scenarios, but they are reminded of a time when a knight killed a troll, a time when a king killed a dragon, and a time when a knight wounded a dragon and then negotiated with it. These are all three variations of the original specification (dragon becomes troll, knight becomes king, and kill becomes wound). Turner's system makes incremental changes to the initial specification in an attempt to retrieve useful cases from memory.

However, Turner's scheme has two drawbacks. First, it partially relies on a careful placement of concepts into a semantic network memory in order for proper similarity to be noticed. For example, knight and troll are both grouped under *violent characters* so the above scenario could have become *A troll killed a dragon*. Instead of this hand-tailoring of the knowledge in memory, we propose a general semantic memory organization, but allow *function tagging* of concepts. This addition allows the semantic memory to appear to dynamically reconfigure itself based on functionality. Thus, a knight and a king are “closer” in memory if the reasoner is attempting to retrieve a nobleman. On the other hand, if the reasoner is thinking about the knight as a fighting agent, then knight and troll would be “close” in memory. The second shortcoming of Turner’s model of creativity comes from the same place as its power. He makes use of a set of 24 heuristics which guide the creativity process, such as *change the agent involved to one of similar nature*. The incremental changes give a great deal of power to the system, but Turner never satisfactorily explains where the heuristics come from or why he settled on the number that he did.

We use the term *imaginative memory* in a slightly looser way than in Turner’s work. A reasoner can request a memory retrieval based on a set of index cues. If all the cues match some item in memory, an exact retrieval is possible. However, the memory system should also return near matches. Some of these near matches will be fairly mundane variations of the exact concept being probed for; others will be more imaginative. It is important to realize that the reasoner has no control over this phase of the CUP algorithm; the memory system is acting at a level too low for the reasoner to be aware of. In functional terms, this simply means that the metacontrol supertask does not have access to that aspect of the memory supertask’s operation. In Turner’s research, the distinction between what a reasoner has control over and what they did not was never clearly defined or explained. Finally, Turner’s approach uses an iterative process which continually makes small changes to the retrieval request; this gradually increases the range of possible retrievals. The memory supertask in our theory has been somewhat restricted in this area; the supertask cannot perform as many iterations during a memory retrieval. This is due to the fact that imaginative memory is only the first step in the entire CUP algorithm; as a result, our view of the technique is inherently more limited than Turner’s. The power that the technique loses at this stage is made up by the later steps of the CUP algorithm.

6.5.2 Analogy

The second step of the CUP algorithm involves what to do with the items retrieved from memory. If the retrieved concept is not an exact match, this is the signal that normal understanding has failed. The memory system had some matching criteria that determined that the retrieved concept is a potentially good candidate with which to understand the new concept. The reasoner must now attempt to determine why this is the case. To do this, analogy is used. *Analogy* is the process by which two concepts are seen to be related in some way (e.g., Falkenhainer, 1987; Gentner, 1989). A well-known analogy is that of the atomic structure being analogous to the Solar System. In analogy terms, the object which a reasoner is attempting to explain is called the *target* and the known analogue is the *base*; the reasoner discovers the relationships which exist between base and target. In the atom example, the atom is the target and the solar system is the base. This framework can be the springboard on which creative ideas are based. If a reasoner can find and make use of an appropriate analogue, then the resulting behavior and/or product will at least appear to be externally creative.

6.5.3 Baseless analogy

Step three of CUP concerns what the reasoner does if straightforward analogy fails to glean an understanding of the considered concept. If this is the case, baseless analogy is employed. Humans are more flexible than the above description would lead one to believe. In fact, humans seem to be able to perform a type of analogy in which there is no base. *Baseless analogy* occurs when the reasoner dynamically builds a base in an attempt to understand a concept. The target in question exists within a specific domain; if the reasoner possesses a great deal of knowledge about some other domain and a shallow amount of knowledge about the target, it might be possible to shift the target from its original domain into the well-known one. Doing this, the reasoner creates a base by relying on their knowledge of the new domain. This form of analogy can most easily be seen when a concept is being explained to someone. The teacher may explain the concept by appealing to the background information possessed by the student. For example, atomic structure can be explained within the framework of gang warfare.[3] Baseless analogy relies on the reasoner possessing a great deal of background knowledge about

---

[3] There was once a large area of the City which was a desolate wasteland of condemned buildings, known simply as the Neighborhood. The New Boys, a street gang, currently control the Neighborhood and have their headquarters in the exact center of the wasteland. Lately,

the domain onto which an analogy is being attempted.

6.5.4 Problem reformulation

Finally, if all of the above steps have failed to produce a satisfactory level of understanding, the reasoner must resort to *problem reformulation*. There are some cases in which the initial statement of a problem or task description is not the one which will lead to an optimal solution. By recasting the problem in a new way, a reasoner may gain a previously unseen insight into the resolution of the situation. Recent work on this topic has been done by Jones (1992). His thesis is that the reformulation should be from an actual problem to a more abstract description; in particular, a primary source of culturally shared abstract knowledge exists in the form of proverbs which reasoners can make use of in solving novel problems. We view problem reformulation as generally being the transformation of one instantiation of a problem into another one, rather than the transformation of a problem into a more abstract form.

Problem reformulation is also the process by which creative reverse engineering occurs in the CUP algorithm. By successively reformulating the given problem description, it is possible to reverse engineer a chain of creative reasoning which would lead from a mundane object to the creative artifact in question. Consider a reader attempting to understand a unicorn who was totally unfamiliar with the concept. If the problem becomes *understand an animal which is like a unicorn but has no horn* (the removal of an attribute), the next cycle through CUP will retrieve the concept of *horse* from memory. The reasoner has reverse engineered *unicorn* to *horse* by removing the attribute of *horn*. To understand the unicorn in this way, then, the reasoner need only transform $horse + horn \Rightarrow unicorn$.

6.6 Function-driven morphological synthesis

For creative understanding to occur, four processes must successfully interact. As described above, these are memory retrieval (possibly imaginative), analogy, baseless analogy, and problem reformulation. We must now specify mechanisms for these processes. Imaginative memory and problem reformulation are performed by a common mechanism within our theory. To see how this is accomplished, we begin by considering the converse of baseless analogy, targetless analogy. *Targetless analogy* occurs when a reasoner begins with a known concept and tweaks it in various ways to discover new concepts which were not previously known. Taking the concept of a horse, for example, and adding wings would result in a pegasus. The mechanism by which this tweaking is accomplished is through *constraint manipulation*. The earlier mechanisms of attribute listing and morphological synthesis also can be described as techniques which perform constraint manipulation. The limitations of both were mainly caused by the enforced restriction that the techniques only alter primary attributes and constraints. In one sense, this is not a bad heuristic; altering secondary characteristics is likely to lead to only A-novel creations. On the other hand, it does restrict the techniques to the formation of E-novel items; R-novelty is unlikely (if not totally impossible).

In an attempt to overcome this serious shortcoming, we have developed a new mechanism, known as *function-directed morphological synthesis* (FMS), depicted in Figure 5. It is assumed that the reasoner has an artifact which needs to be understood (either a goal (G), a solution (S), or a critique (K)). The reasoner applies a set of manipulator functions to the artifact, altering its attributes and producing new artifacts. If only the artifacts which possess the original functionality and are novel to the reasoner are considered, creative artifacts are the result.

---

a new gang calling itself the Elected Ones has shown an interest in acquiring control of the Neighborhood. Fearing the loss of control, the New Boys decide that they need to bring in an outside gang to aid their cause. The New Boys learn that the blood enemies of the Elected Ones are a group of men known simply as the Pros. The Elected Ones and Pros so fear each other that as long as one Pro is in the New Boys' territory for each Elected One there will be no violence. So, the New Boys bring in the exact number of Pros as there are Elected Ones. Frustrated in their takeover attempts, the Elected Ones are forced to stay outside the Neighborhood and only circle it. After a while, all three gangs decide to change their names to include "-tron" in their names. So, the Elected Ones become the Elec-trons, the Pros become the Pro-tons, and the New Boys become the New-trons. The Elec-trons continually circle a vast area of open space. At the center is a group of Pro-tons and New-trons. Notice that there is always the same number of Pro-tons as there are Elec-trons. (WKRP Episode 60, 1980)

- Consider an artifact in the world, which will be designated as M. For creative understanding, M will be either the goal (G), the solution (S), or the critique (K).
- Let f be defined as the function which returns the function of an artifact.
- Let C be the class of functions which alter an object, either by changing some attribute of that object or by adding a new attribute to the object.
- C_1 through C_n are a set of n such functions.
- Thus, a set of objects S_{all} can be created by $\cup_{i=1\ldots n} C_i$.
- Consider the subset, S_f defined as $\{s | s \epsilon S_{all}$ and $f(s) = f(M)\}$
- Finally, consider the subset S_c defined as $\{s | s \epsilon S_f$ where s is unknown to the reasoner $\}$
- The items in S_c are useful (they fulfill the same role as the original object M) and they are novel to the reasoner. Therefore, they are creative.

Figure 5: Function-driven Morphological Synthesis

FMS can exist in both a strong form and in a weak form. Strong-FMS performs the changes to an object by examining other objects with the same functionality to see how they accomplish their tasks. Weak-FMS does away with this constraint and guides the manipulation through the reasoner's knowledge of a given attribute and its possible values. There is still the problem with constraint manipulation in the sense of exactly what is allowable. This issue will be touched on in Section 7.3 when we discuss our knowledge organization scheme.

Returning to our luggage example of Section 5, let M = *piece of luggage*. Then, $f(M)$ = *act as a closet/dresser for a traveler while away from home*. The types of alterations which were discussed earlier will still be allowed. In addition, the types of alterations which attribute listing and morphological synthesis were incapable of performing will now be possible.

If the memory supertask uses FMS to alter a probe for a memory search, the observed behavior will be akin to that of imaginative memory. If FMS is permitted to act on the problem description given to the understander, then the observed behavior will be that of problem reformulation. It is this power which allows FMS to be the driving force in the entire creative understanding process.

7 When to say when: How far to manipulate constraints

A central issue in constraint manipulation systems is how does a reasoner know exactly which manipulations are "good" ones and which ones are potential dangerous (or just useless). Various approaches have been taken to attempt to minimize this problem in other systems which attempted to be *creative* by some metric. One end of the spectrum is to have the system do little self-monitoring; this approach can be seen in Lenat's AM system (Lenat, 1990). AM was a system which discovered new mathematical ideas through a process of exploration and constraint manipulation. It was unable to do much filtering, however, as Lenat himself would sift through the output to point out to AM what concepts were worth exploring. So, while it did create a number of useful and perhaps creative concepts (such as prime numbers and Goldbach's conjecture), it also created a much larger number of worthless concepts which were filtered out by the human researcher. Towards the other end of the spectrum, we have systems which had built-in heuristics for deciding what to manipulate and how far to carry the alterations. Turner's MINSTREL system is one such example; SWALE (Kass et al., 1986) is another well known one. Both could be considered to have done some degree of creative reasoning by appealing to these human-supplied control heuristics. While this is certainly a good approach (especially in the initial stages of research), we would prefer to develop a technique which would allow for more flexible restraint to be exercised, decided (at least in part) by the system itself. We begin by examining what kinds of constraint modifications can take place within a reasoning system.

7.1 Types of constraint modification

There are two ways in which constraint modification may occur within the framework of FMS. The reader can relax the definitional constraints which exist on the concept. For example, a horse's primary mode of locomotion is legs. Relaxing this restriction, the system may decide to consider a "horse" whose primary locomotion is wings, resulting in a pegasus. The other option is to add a new aspect to the concept under consideration. A suitcase, for instance, would not have any information regarding a mode of location. By adding this, it would be possible to create an independently mobile suitcase, much like the one depicted in Terry Pratchett's *The Colour of Magic* (1983).

7.2 Problems with constraint modification

The problem with undirected constraint manipulation is that the reader has no way of knowing how far to allow the alterations to occur. In the above example, toaster could have been made the means of horse locomotion. Or a reader might decide that love is a good method of transportation. Up to a certain limit, constraint manipulation will result in concepts which could be called creative. After this point, the concepts are too bizarre to be generally useful. A "bizarreness" threshold exists: we like and understand things up to a certain level of bizarre attributes; after that, we begin to feel uncomfortable with the concept and reject it. It is possible for the programmer to enforce bounds on the level of bizarreness allowed to occur, but this seems to restrict the outcomes in an unreasonable fashion (what if the story actually did involve a toaster-horse?). Alternatively, the system could simply make arbitrary relaxations. Both of these options have been argued against from a theoretical perspective (Birnbaum, 1986). We prefer to have reasonable relaxations emerge from a solid theory of knowledge organization and from enforcing the idea that relaxations must be directed towards an ultimate goal.

7.3 Knowledge organization scheme

Our knowledge organization scheme resembles a standard semantic network, but knowledge is tagged through the use of a multidimensional grid, as shown in Figure 6. One axis of the grid represents a Schankian breakdown of knowledge: *action*, *agent*, *state*, and *object* (Schank & Abelson, 1977). The other dimension represents a more-or-less "natural" breakdown: *physical*, *mental*, *social*, *emotional*, and *temporal*. For example, a TRANSfer is a generic action. In the physical column is the PTRANS of Schank's work, the mental column contains MTRANS, and the social column contains ATRANS. Our extended representation also includes emotional TRANSfers, like the giving of one's love to another; and temporal TRANSfers, such as the act of March getting closer to us in time.

The other difference between a standard semantic network and our knowledge organization scheme is the *function tagging* of each concept in memory. Each concept within our knowledge system is tagged with the set of common functions which it can be used to perform. If the reasoner is searching for similar concepts to a given one, this search can be aided by the presence of these function tags. In one search through memory, a car and a horse would be discovered to be closely similar; with a different function in mind, a horse and a zebra would be discovered to be close in the semantic net. This tagging allows a more flexible organization of knowledge with the semantic network than previous methods which were forced to employ a careful placement of items into the network to ensure that the proper similarities were revealed when needed.

Coupling our knowledge organization scheme with the function-driven morphological synthesis discussed above results in constraint relaxation and manipulation which is bounded in a reasonable fashion. Each change on a concept from the FMS may leave the concept in the same conceptual grid cell (an *intracellular shift*) or it may cause the concept to cross a cell boundary (an *extracellular shift*). Additionally, each FMS change costs the system. So, a conceptual shift which consists of only two FMS steps is easier to perform than one consisting of six FMS steps. A cell boundary crossing costs significantly. As a result, conceptual movement within the same grid cell is the cheapest type to perform. Movement along *either* a row or a column is more difficult, and movement which must go along both is the most difficult to perform. These costs act as a heuristic which guides a reasoner in performing creative understanding. The greater the cost, the more conceptual movement has

	Physical	Mental	Social	Emotional	Temporal
Agents	person	consciousness	boss	Ares	entropy
Actions	walking	thinking	selling	loving	getting closer to March
Objects	rock	idea	teacher-student relationship	hatred	second
States	young	lack of knowledge	public dishoner	being angry	early

Figure 6: Knowledge representation grid

occurred. As a result, the higher the cost, the more likely it is that the result will be more bizarre than creative. This knowledge organization theory allows for constraint relaxation to take place in a reasonable manner.

Within the creative understanding process described by the CUP algorithm, this cost of movement influences how many iterations are permissible. Each successive cycle through the function creates concepts which are more and more distant from the original one. The first few iterations will result in concepts which fall within the same grid cell as the original concept. More cycles will create concepts which are shifted in the grid with respect to the beginning concept. By tracking this movement, the reasoner can decide when creative understanding has become too expensive to allow to continue, based on the goals and tasks of the reasoner.

8 Implementation

The ISAAC system is currently implemented at a level of functionality capable of reading the story *Men Are Different* (Bloch, 1963). ISAAC is built in Common Lisp and uses the KR frame package (Giuse, 1990) for basic knowledge representation. ISAAC currently uses the COMPERE parsing system (Mahesh, 1993) as a drop-in module for its sentence processing supertask. The union of the two systems is not yet seamless, so ISAAC's higher-level reasoning actually processes a mixture of representations produced by the COMPERE parser and representations produced by hand-translation of certain "difficult" sentences (the ratio of COMPERE sentences to hand translated ones is approximately 3 to 1). At present, ISAAC is capable of building a detailed story structure model, a scenario model of the events depicted, and a metareasoning model of its own activities during the reading process. On the knowledge side, there are currently a few hundred concepts stored within the system. This knowledge is organized according to our knowledge organization grid.

The story we are currently working with is the example one presented in Section 2.2, *Men Are Different* by Alan Bloch (1963). ISAAC processes the understanding of this story in the following ways. ISAAC is initially told to read the text file containing *Men Are Different*.[4] ISAAC is expressly told that the file it is to read is a narrative science fiction text; as a result, ISAAC's metacontrol decides that the first thing it needs to accomplish

---

[4] *Read* in this case means to consult a disk file line by line in order to simulate the reading process. While we do not expressly model eye-movement, as in Just and Carpenter (1988), ISAAC does have the capability to skip around in the file in a nonlinear fashion. This approximates a human reader's ability to reread sections of the text, if needed.

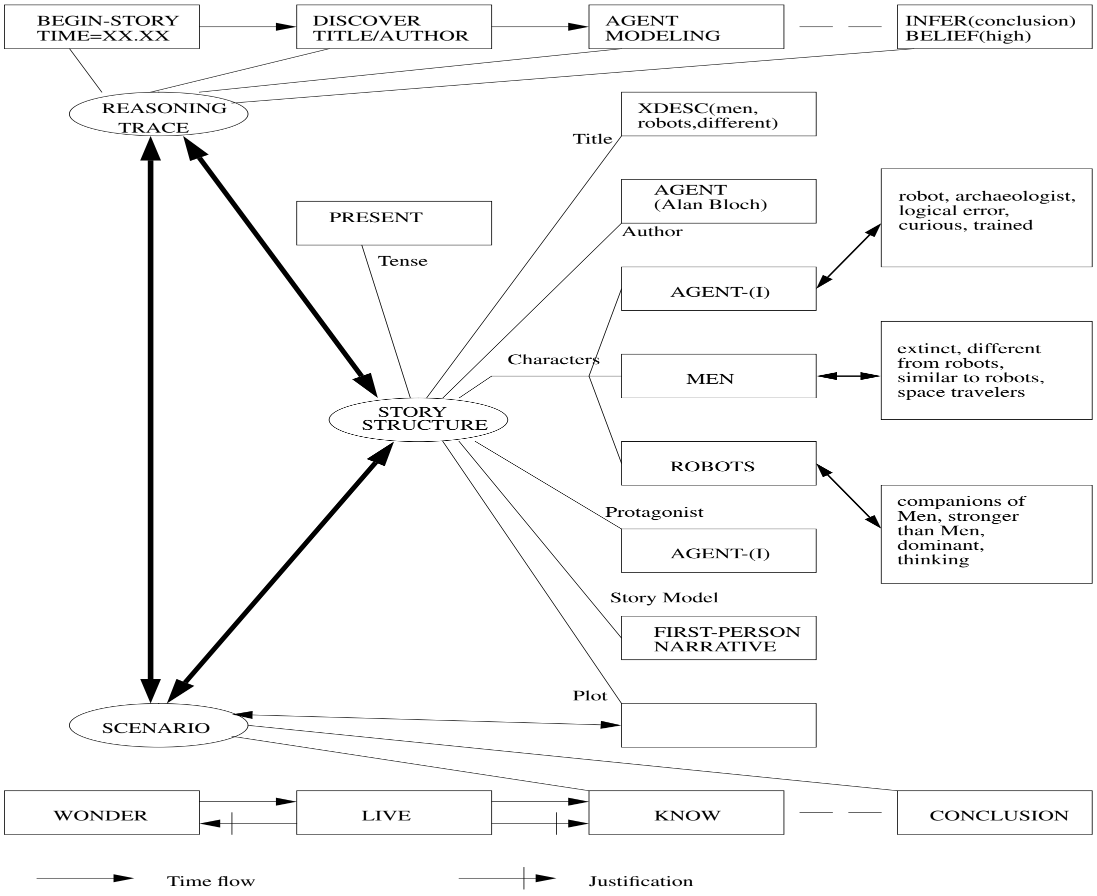

Figure 7: Partial view of ISAAC's understanding of *Men Are Different*

is to see if the text has an attached title and/or author. ISAAC's generic narrative story structure model includes the information that a set of text with a phrase and an agent prior to the beginning of the first paragraph can be assumed to have an attached title and author. Metacontrol makes use of this, treating the stand-alone phrase *Men Are Different* as the assumed title, and the supplied agent *Alan Bloch* as the assumed author. The author requires no further understanding, it is simply recorded as the author of the current story and as being attached to an agent in the same reality that ISAAC is in. The title is passed on to the other supertasks so that further processing of it may occur. The sentence processor parses the sentence for ISAAC; since ISAAC is starting the reading experience, it is unable to provide any hints to that supertask. The phrase is recognized as an extended state declaration: ISAAC expects an object to be related to another object in a given way. Scenario understanding tags *Men* (a type of agent which ISAAC is familiar with) as the first object and *Different* is as the relation in question. Metacontrol now expects to find the second object; namely, what *Men* are *Different* from. At this point, no such answer is available. Metacontrol decides, due to its knowledge of the importance of titles in most texts, that this is an important omission. ISAAC expects to find the answer to this pending question in later text pieces. Now that the title and author have been understood, various memory requests are generated. First, the title and author are remembered (more specifically, the memory supertask is requested by metacontrol to store the information). Next, a memory probe is requested to see if the system has either read this story before, read other Alan Bloch works, or read other stories about men being different. Since this is the first story episode, nothing is returned.

ISAAC is now ready to begin reading the body of the story. The *pronoun reference* task is responsible for determining that the *I* in the story refers to an agent whose occupation is archaeology. At this point, the story structure model being used to guide ISAAC needs to be changed. The generic story ISAAC knows about is a third-person narrative. The sentence processor has just found a pronoun acting in the role of an agent. This indicates to the genre identifier that it cannot be a third-person narrative, since it is known (also from the sentence processor) that the sentence in question is simple declarative. ISAAC does a memory retrieval and discovers that first-person narrative more closely resembles the story thus far and switches to it. This gives the story structure supertask the protagonist of *Men Are Different* for "free." *Punctuation analysis* (a part of the sentence processing supertask not covered by COMPERE) notices that *Men* is capitalized in a way which does not follow grammar rules, implying that the word carries special meaning. ISAAC's metacontrol decides to remember this anomaly as potentially important. Meanwhile, *character identification* adds the fact that *I* is a story character. By the end of the first paragraph, ISAAC realizes that something is not quite right. ISAAC knows about modern-day robots which are industrial machines, capable of performing autonomously in various highly repetitive tasks, such as spot welding on a car. The robot which is being described in the first paragraph is very different from this. ISAAC's metacontrol decides that it needs to attempt to creatively understand the concept of robot it is being presented with and suspends its disbelief in such an object for the benefit of the story. This decision is motivated by ISAAC's inability to understand the story within the conceptual framework initially possessed by ISAAC.

Creative understanding is now faced with the task of explaining the robot that ISAAC sees in the story. A normal memory retrieval returned a concept of robot which is insufficient to explain the story character. Concepts which were imaginatively retrieved from memory (factory machines, mechanical tools) are also too limited to provide an adequate explanation. Creative understanding now attempts baseless analogy on the concept of robot. Taking robot as an industrial tool, ISAAC attempts to transfer the concept to the domain of volitional agents (since this is how the story robot appears to be acting). This attempt is somewhat less than totally successful; in this case, ISAAC does not have enough background knowledge to allow the success of baseless analogy in this case. Next, ISAAC attempts problem reformulation, using FMS to produce a merged concept containing elements of the current robot concept and the best volitional agent it can retrieve from memory—a man. The result is a man-like (in form), intelligent, volitional robot. Man-like is used since ISAAC has now answered the question of what *Men* are *Different* from—robots. ISAAC's understanding of standard narrative structures is that authors do not waste time presenting the reader with the blatantly obvious. So, in addition to being somehow different from robots, men must also have some similarities to robot (after all, it does not make sense to say *Men are different from three ton boulders of granite*; it's an obvious fact). As such, the new concepts maintains the characteristics of the original robot in the sense that it is made of metal, resistant to damage, uses sensors and feedback as a control mechanism, and so on. This concept is sufficient to understand the robot in the story. ISAAC then store the new concept in memory as a story-robot.

The rest of the story proceeds in a similar fashion; we will touch on only a few highlights until the final two paragraphs. The main differences to the above procedure occur in the third paragraph. The second paragraph provides additional background, as does the fourth. The third presents the most problem, since it contains a high degree of metaphor. Men do not have bright flames as far as ISAAC knows. At this stage in ISAAC's development, we are choosing to "punt" on this issue; ISAAC's imaginative memory returns a scenario where *bright flame* is used as a description of superiority. This allows straightforward understanding of the sentence. In later versions, ISAAC will be required to attempt creative understanding on this concept as well. *Divine* is a known concept and reinforces the assumptions made about *bright flame*. The sentence discussing the *strands of the web* is left non-understood. We do not provide a direct meaning, and ISAAC is currently lacking the necessary information to successfully understand it. The idea that the current civilization has lost something that Mankind had is understood; the assumption is that it is not a physical object, but something more abstract. When the story understanding is complete, ISAAC is left with an outstanding goal of understanding this concept.

The last two paragraphs act as a "story-within-a-story." The *metacontrol supertask* must enable the system to understand the ending paragraphs in this way. This story, while short, contains a complex temporal structure. Time does not flow linearly in the story; any reader attempting to understand it, ISAAC included, must be able to handle this as it is these last two paragraphs which makes the point of the story. In order to appreciate the ending, the reader must have maintained an accurate model of the narrator, including its beliefs and goals, in order to understand why the robot's mistake took place. ISAAC needs to perform a similar creative understanding task on the man in the story as it did on the robot narrator. In this case, however, the need to perform the creative understanding arises from a need to understand why the robot narrator acted in the fashion that it did. As a result, ISAAC understands that the robot is seeing the man as more similar to itself than is warranted, this explains the logical error.

The key irony in the story can be seen as a dual shift within our knowledge organization grid. ISAAC is presented with a robot character which is acting as an *agent* rather than as a *physical object*. This is the first shift which must take place for ISAAC to "make sense" of the story. The twist ending occurs because the robot narrator decides to treat the Man, a *physical agent*, as a *physical object* and disassemble him with the intent of doing a field repair.

During the reading experience, multiple cases are being built. A *case* is a memory of a specific experience in the life of a reasoner (Kolodner, 1993). First, there is the case representing the events of the story, a summary of what ISAAC understands of the text. Second, this case contains a subcase representing the flashback scene. Since it stands on its own as a separate event, it is an independent case. Third, there is a case which represents the actions of ISAAC as it read and understood the story. This case includes the information as to when the system realized that the main character was a robot and not a man, as well as other key points in the course of the reading experience. It is a trace of what ISAAC went through during comprehension rather than simply being the result of the comprehension task. If ISAAC had greatly empathized with the robotic narrator (currently beyond its abilities), a fourth case may also exist. This would be a sort of "pseudo-experience" in the history of the system—it would remember it as being in the context of a story, but the events will seem more real than that. Such a case in memory could, for instance, aid the ISAAC if it ever has to deal with a person from a radically different culture and it remembers that one should not project a way of life onto another due to surface similarities.

Finally, if ISAAC is familiar with science fiction stories or with this author, the start of this reading experience will trigger remindings which aid the process. ISAAC is familiar with science fiction from the information we have provided it as background. As a result, the idea of a robot narrator will not require as much disbelief suspension as someone who had never read a science fiction story. Recall that for ISAAC to comprehend the story, it must be willing to accept that robots are the dominant lifeform in the future, that humans have practically died out, and that robots are capable of making the logic errors that the narrator did. Had ISAAC refused to accept this world, the story will not be fully appreciated.

A portion of ISAAC's understanding of this story can be seen in Figure 7. The process of reading transforms a text from the given form to a series of representative cases and the three understanding models shown in the figure.

9 Evaluation

The final facet of our theory which we need to consider is that of *evaluation*. The theory needs to be evaluated both as a reading theory and as a creativity theory.

9.1 Reading evaluation

Traditionally, AI reading systems have been evaluated in one of two ways—either by having the system answer researcher-supplied questions regarding the read material or by having it produce a summary of the text (this is true of almost all AI reading research; see, e.g., Wilensky, 1981; Cullingford, 1981; Dyer, 1983; Lehnert, 1977). In either case, the system's output was judged by the researcher. If the question answering approach was used, the questions tended to be of two types: literal questions whose answers were directly in the text, and inference questions which required that the system know more background knowledge than strictly in the text. Summaries

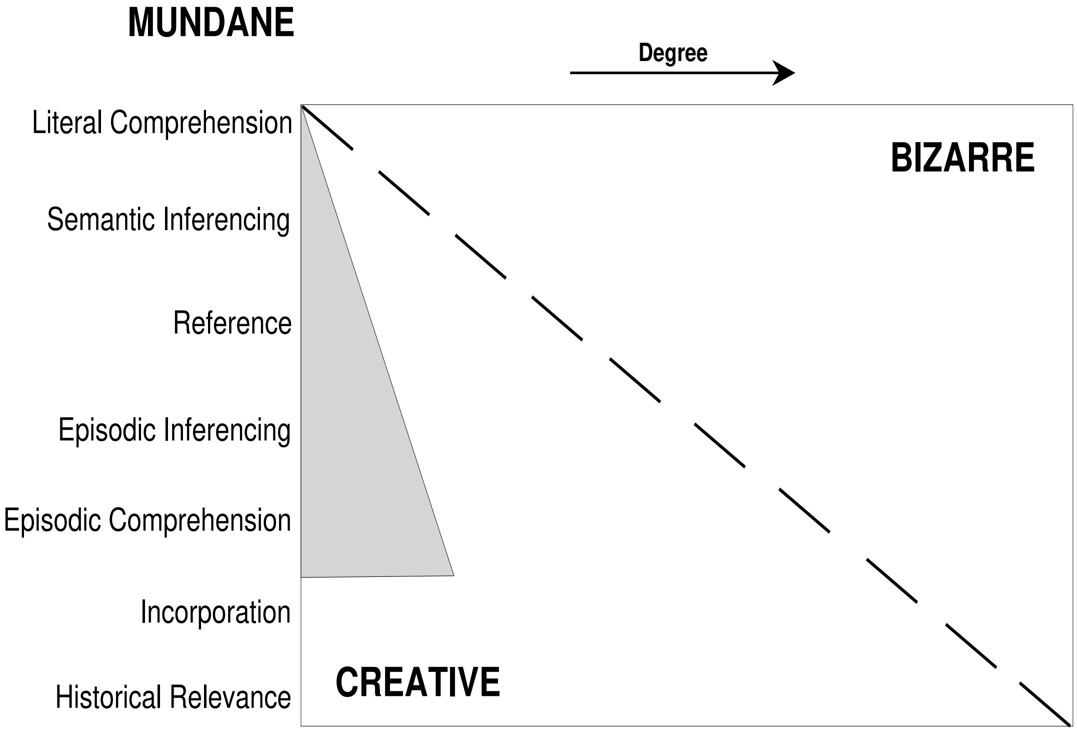

Figure 8: Evaluation

were generally required to capture the main points of the text in a succinct manner. Thus, the AI system was "tested" in a way akin to that of a student being tested for reading comprehension.

With ISAAC, one must be wary of both of these methods. It has been argued that ISAAC will be impossible to judge correctly due to the very fact that it is creative. The argument says that any answer ISAAC gives will be judged correct as it is possible within the bounds of creativity. Consider the following short, short narrative:

> John hit Mary. Mary ran into the house, crying. A minute later, Mary's mother came out of the house and scolded John.

If asked the question *Why did John hit Mary?*, ISAAC might very well respond with *Because space aliens from the Vega system controlled his mind and forced him to do it*. As argued above, there is a point where creativity breaks down and becomes simply bizarre. With no other context, the above answer is probably more bizarre than creative.

To control this facet of ISAAC's evaluation, we developed the comprehension grid depicted in Figure 8. The shaded area represents the scope of previous story understanding systems. The dashed line symbolizes a rough boundary between the creative and the bizarre. It is dashed because the exact location of this boundary will differ from genre to genre and possibly from story to story. Each level lower on the left-hand column represents a type of comprehension which is progressively more involved. Each such level also allows more creativity to manifest without crossing the bizarreness threshold.

Literal comprehension is the simplest form of comprehension. It involves understanding a text sufficiently to answer direct questions concerning it. In the John and Mary story above, a level of literal comprehension would be adequate to answer questions such as *Who hit Mary?* (*John*) and *Who ran into the house?* (*Mary*). *Semantic inferencing* refers to a level of comprehension which adds conceptual knowledge to the literal level. It would allow the reader to handle questions such as *What sex is Mary?* (*female*) and *Did John touch Mary?*

(*yes*). These answers are not directly in the story but can be inferred from the facts that *Mary* is generally a girl's name and that in order to *hit* someone, you have to *touch* them. The next level comprehension, *reference*, involves a more sophisticated understanding of story elements. It would allow the reader to see that *Mary* was *crying*, *Mary's mother* did the *scolding*, and that *the mother* came out of the same *house* that Mary ran into. Next is *episodic inferencing*, which allows more in-depth reasoning about story elements. This reasoning does not result from semantic knowledge concerning story concepts, but from the reader's own experiences dealing with similar scenarios. An example question and answer would be *Why did John hit Mary?* (*She was teasing him*). This response is not a result of the meaning of *to hit* but is plausible given an average reader's background knowledge. *Episodic comprehension* is the next comprehension level. This level represents the reader understanding the entire story as a cohesive set of events and seeing the underlying causal patterns. At this point in the comprehension scale, the reader would be able to supply a title for the story (in this case, perhaps *The Quarrel*) and to answer more abstract questions, such as *Should Mary's mother have scolded John?* and *What ages do you think the characters are?*.

The above levels of comprehension represent roughly what previous story understanding systems were capable of accomplishing. The remaining levels represent higher comprehension levels which human readers possess. Episodic comprehension can be carried a step further to arrive at the level of *incorporation*. At this level, the reader understands the story as a cohesive set of events, can see how the story relates to other experiences in their own life, and can use the experience for reasoning about future events which are similar. There is one final level of comprehension, *historical relevance*. This is the most elusive of the levels and requires the most sophistication on the part of the reader. Some stories' comprehension rely on the reader understanding the historical period in which they were written. For example, Orwell's *1984* (1949) is better understood if the reader is aware of the sociopolitical climate of the late 1940s, in which it was written.

Reading teachers and researchers have broken comprehension down in an analogous way to our seven levels. Teaching reading textbooks (e.g., Strang et al., 1967) describe three levels of comprehension. The *literal comprehension level* is when the reader understands the literal meaning of the text. This level is represented in our breakdown by the first five levels and makes up the largest portion of reading comprehension. The next level, *reading between the lines*, is when the reader can accurately reason about the intents and purposes of the author of the text. Characters are identified with, truth value can be assigned, various points of view may be recognized, and the reader is capable of understanding metaphor and irony. In our seven-level comprehension table, levels four, five, and six represent this level of understanding. Finally, there is the level of *reading beyond the lines*. The reader fully incorporates the read material into their own background, learns new concepts from it, can place it within historical relevance, and has a complete understanding of the text. Levels six and seven of our breakdown correspond to this final level of understanding.

To be judged a success, ISAAC must be capable of answering questions and generating summaries which fall within that region covered by previous systems. In addition, ISAAC must demonstrate some level of comprehension beyond these earlier systems in order to be viewed as an improvement. ISAAC's creativity will be judged by how much of the creative portion of the grid it manages to cover without degenerating into the bizarre zone. This coverage will be judged based on the answers to questions posed to the ISAAC system and on the summaries it produces. Since we are using real stories, we have an opportunity to make use of real evaluation questions, as well. Turner (1992) appealed to everyday people to judge the level of performance of his story generation system, MINSTREL. We intend to follow a similar path of removing some of the evaluation from the researchers' control. To achieve that end, we plan to ask high school English teachers to read the stories we provide and generate questions covering the material. After ISAAC has been "quizzed" on the stories, the answer set will then be graded by the teachers and a grade assigned. While only one metric of evaluation, we feel it will be an important one.

9.2 Creativity evaluation

If ISAAC is to be accepted as a true "creative" reader, then there must also be a way of evaluating its creative potential. This is more difficult than reading evaluation; while educators do believe they have fairly accurate ways of judging reading ability, no one claims to have a similar degree of success with judging creative acts.

Internal	External	Concept production	Concept combination	Result
NO	NO	NO	NO	Mundane result
NO	YES	NO	YES	Other combinatorially creative
NO	YES	YES	NO	Other productively creative
NO	YES	YES	YES	Other creative
YES	NO	NO	YES	Self combinatorially creative
YES	NO	YES	NO	Self productively creative
YES	NO	YES	YES	Self creative
YES	YES	NO	YES	Significantly combinatorially creative
YES	YES	YES	NO	Significantly productively creative
YES	YES	YES	YES	Significantly creative

Figure 9: Views of creativity

Part of the problem, as discussed earlier, is that creativity assessment is a highly subjective thing.

It is almost certain that ISAAC will act in a way which it views as creative, in much the same fashion as a child discovering new concepts. As with the child, it is also likely that much of what ISAAC comes up with will be well known to the researchers involved. It is possible to extend our earlier internal/external creativity description in an attempt to get at a better understanding of how a creative program would be judged. As described by Leary (as reported in Taylor, 1988), it is possible to break creativity into the production of new concepts and the combination of existing concepts. Doing so, results in four types of individuals being identified (the following also comes from Taylor, 1988):

- *reproductive blocked*: This person is mundane, no new concepts are created and there are no new combinations of existing concepts.
- *reproductive creator*: No new concepts are directly generated, but this individual is capable of creatively combining existing concepts.
- *creative creator*: The creative pinnacle, this person creates new concepts and can successful use them in new ways.
- *creative blocked*: The final category includes those individuals who are capable of producing new concepts but is unable (or possibly unwilling) to use them in any way other than a in a known, mundane fashion.

Combining these four creativity types with our earlier discussion of internal/external creativity, we get the table depicted in Figure 9. Notice that all of the possible sixteen combinations are not represented; some of the possible ones have no real meaning. Figure 9 is not intended to indicate that those ten divisions are the only ones which exist in a creativity spectrum. The creativity spectrum extends infinitely from the beginning point of mundane to the more and more creative; the ten divisions identified in the figure represent “signposts” along this infinite pathway. The average creative act, whether invention or understanding, will contain elements of several of these points. The ten degrees of creativity identified in Figure 9 would map onto the vertical axis of Figure 8. This does not mean that we (or anyone) can arbitrarily decide at what point the comprehension at a given level crosses over from the creative to the bizarre; the scale is not the same for each level of comprehension. On the other hand, the creativity scale, as rough as it is, does give some objective evaluation potential to the process of judging a so-called creative episode from a reasoner.

9.3 Genre notes

Finally, a note must be presented concerning the nature of the genre of stories we have chosen for ISAAC. As already mentioned, science fiction stories offer tremendous opportunities for creative understanding; indeed,

waterbed
waldoes
communication satellites
the term *robotics*
rocket packs
personal communicators
diagnostic medical beds
hypospray injectors
hand-held calculators
ring-activated guns
CAD/CAE/CAM

(a) Science fiction ⇒ science fact

space elevators
light sails
terraforming
transporters
faster-than-light travel
time travel
Dyson spheres

(b) Still science fiction

Figure 10: Science fiction concepts

many science fiction works force the reader into a position of needing to perform creative understanding, especially if they are a novice science fiction reader. Many of the devices described in science fiction stories have predated actual real world artifacts of the same type. For example, the artifacts in Figure 10(a) all existed in stories prior to existing in reality. The items in Figure 10(b) represent a handful of artifacts which currently have no real world instantiation. It should be clear, then, that science fiction offers countless opportunities and requirements for creative understanding to be performed.

There is also an inherent danger with short stories in general which we must guard against in our selection process. Since the mechanism of communication is limited to a few hundred words, authors may count on the reader's background filling in crucial bits of information. This is perfectly acceptable unless the needed bit of information is extremely story-specific. If this is the case, then ISAAC would seem to be too tailored to the particular set of stories given it. We have tried to avoid this pitfall by choosing stories which, while requiring significant amounts of creative understanding, hinge on more-or-less general knowledge (at least across the domain of science fiction stories). *Men Are Different*, for example, embodies concepts of the future, space travel, intelligent machines, the extinction of humankind, and so forth; all of these concepts may be useful in other science fiction stories. In contrast, there is the short story *Blood* by Fredric Brown (1992). It involves the time traveling of two vampires, who are fleeing into the future. They hope to find a future where knowledge of vampires has been lost and they may feed in peace. After a few narrow escapes, they arrive in the far future to be confronted by a telepathic race of new creatures (humans have long since died out). The creature confirms that it knows nothing of vampires, this excites the protagonists greatly. Unfortunately, the final line of the story has the lifeform revealing that it is actually an intelligent, evolved form of the turnip. Naturally, the saying that *You can't get blood from a turnip* must be known to fully appreciate the story; while this knowledge could be given to ISAAC, it seems too specific to be of use in a wide variety of stories.

10 A critique of past reading research

A methodological goal of any complete theory should be the formation of a common framework in which past theories and systems can be placed. Our functional theory allows us to judge past attempts from several perspectives. The most important ones are:

- A critical examination of past systems in order to discover which tasks and supertasks they attempted to cover.
- A similar examination carried out on past theories which do not exist in a system form. This includes a critique of the processes and knowledge that each identified as important or central.

- A critical review of systems with regards to the types of knowledge each understood.
- An examination of the creative understanding aspects of past reading systems.

10.1 A critique of past reading systems from a task perspective

Past reading systems have not been greatly successful due to their insufficient coverage of the set of tasks which reading involves. Early examples such as PAM (Wilensky, 1981) and SAM (Cullingford, 1981) were narrow attempts to handle scenario understanding through plans and scripts. They did not try to benefit from other aspects of the reading process; furthermore, they dealt with extremely short narratives. Dyer's BORIS (1983) was a more involved system which attempted to overcome the deficiencies of earlier ones by integrating the current theories of its day in order to perform "in-depth" reading. Unfortunately, it too overlooked some crucial aspects of the reading process which we have identified, such as the story structure. BORIS also ignored the human facility to skim and skip portions of a text and attempted to read everything in as much depth as possible. This meant that a huge amount of internal representation was required before reading could take place, placing a limit on the number of stories which could be read. Later systems tried to vary the reading depth, such as IPP (Lebowitz, 1983) and AQUA (Ram, 1991). Unfortunately, IPP failed to do any high-level reasoning, such as explanation of anomalies. AQUA added the explanation and reasoning supertask, as well as some of the memory supertask, but it ignored the story structure aspects which could have been a valuable aid.

There were other systems being developed which carried scenario understanding to a more refined level by augmenting it with the explanation and reasoning supertask we have identified. These attempts placed a greater emphasis on *plan recognition* than some other systems (such as PAM) had. Levy (1979), for example, believes that this identification of characters goals, beliefs, and plans is the crucial aspect of reading which needs to be studied. Grosz and Sidner (1986) share this assumption; they see reading comprehension as consisting of three parts: linguistic, intentional, and attentional. The linguistic portion handles sentence processing, the intentional manages the goals, plans, and beliefs of the participants in the text, and the attentional handles the way in which the reader relates to the text. This model is similar to our theory but does not make use of an explicit story structure understander, an explicit memory management supertask, or an explicit explanation supertask. As such, it covered more of the reading process than other attempts, but did not attain complete coverage of the tasks which we have identified.

There were other attempts to use the idea of plan recognition as the basis for a reading comprehension system as well. The idea of a text communicating the intentions and plans of the characters involved can be traced back to the philosophical work of Grice (1957). His thesis was that to understand language (or text) completely, one must not only understand what the words mean but what the beliefs and intentions of the speaker (or characters) are. Prior to the work which attempted to make tenable systems based on this, there was a great deal of research in AI dealing with *planning* (for early, foundational work in this area, see Newell & Simon, 1972; Sacerdoti, 1977; Fikes & Nilsson, 1971). By making use of this, some systems attempted to directly understand text by identifying the plans of the characters. Allen (1979; 1983) and Wilensky (1981) were among the first to attempt this. Others included Carberry (1988) and Sidner (1985). Plan recognition is a powerful technique, particularly when the reader is able to understand the interactions of characters and their plans, rather than understanding a single character (Wilensky, 1983). Still, it is too limiting to be a complete model of reading; a system concentrating on just plan recognition falls into the same trap as all the systems described—concentrating on too small an area of the reading process.

Other systems also concentrated on various narrow aspects of the total problem. CYRUS (Kolodner, 1984) concerned itself with the memory management issues, but slacked in the other aspects of reading. Not all story understanding systems overlooked the story structure or the sentence level as an aid. For example, story grammar theory (Rumelhart, 1977) explains stories much like our story structure supertask does—they are coherent text passages with characters, a setting, and a series of actions (the plot description). Unfortunately, story grammars do not go beyond this level; they ignore scenario understanding, sentence processing, and reasoning. There have also been numerous systems which concentrate solely on the sentence processing level. These include ATNs, statistical analysis, and more sophisticated parsing methods. Some had a great deal of success; however, there are many sentence problems which can be more easily solved through the use of higher level processes (Birnbaum, 1986). The sentence processing supertask must be interwoven with the entire system. Without doing this, it is difficult to move much beyond sentence level comprehension.

While they made significant contributions to the areas they focused on, these past systems all ignored aspects of the total reading process. Because of this, they had to work harder at the portions they did consider. Our claim is that the set of reading tasks we have outlined reinforce each other, aiding one another during comprehension. A complete understanding of the tasks and their interaction will enable better reading systems to be developed. Our theory has tried to address past systems' deficiencies by making use of what was learned from each attempt.

10.2 A critique of past theories of reading

The systems described above were all the result of artificial intelligence research with various degrees of influence from psychology. However, the product of AI research is not the system itself, it is the theory behind it. In some way, then, the critiques above were of theories rather than of systems. We concentrated on the systems for the reason that some of the described work was not intended to be an actual reading theory/system; instead, the work used reading as a task to demonstrate the research results. There does, however, exist a long history of research on the question of what reading is, from both the psychology and teaching reading discipline; the line can be traced back to the publication of Huey's influential text, *The Psychology and Pedagogy of Reading* (1908). Since that time, numerous theories of what reading is and how it is accomplished have been proposed. A close examination of these theories reveals that a wide amount of variation exists among them, with several levels of description being concentrated on.

Graesser (1981) described six basic knowledge sources involved with textual comprehension: linguistic, rhetorical, causal, intentional, spatial, and roles, personalities, and objects. Mackworth (1972), on the other hand, took another approach and chose to concentrate on the processing which goes into reading, developing a model of the reading task which was largely an information processing one. Other well-known reading researchers also concentrated more on the process than on the knowledge needed. One of the most detailed models was developed by Gough (1972) who came up with a model of "one second of reading," which was designed to describe all of the processes making up one second of oral reading.

Another level of description which researchers considered was that of the *skills* needed for proper reading comprehension. Davis (1968) identified five skills necessary for reading to occur. These were: memory for word meanings; drawing inferences from the content; following the passage's structure; understanding and/or recognizing the author's purpose, attitude, tone, and mood; and being able to answer questions, either explicitly or in paraphrase. Notice that these skills are sort of an intermediate level between knowledge and process and include elements from both. A similar model was suggested by Clymer (1968) which broke reading into decoding, grasping the author's meaning, testing and recombining the author's messages, and application/extension of the read material. A very extensive view of the make-up of reading was the *Barrett Taxonomy, Cognitive and Affective Dimensions of Reading Comprehension*, also reported in Clymer (1968). A notable addition to past taxonomies was the category of *appreciation* which involved emotional response to the text.

The final two models which we wish to touch on are possibly the most influential to date. The first is the result of extensive research by Kintsch and van Dijk (1978). Their model proposes a unified process which is responsible for all of the characteristics of reading described in the above attempts. Text is transformed into *propositions* (units of information) which are combined by the reader into higher level understanding structures. The model has been extremely successful in modeling psychological data regarding the reading process, particularly those research dealing with recall of read material. The model has also been updated and expanded by a number of researchers, including the original creators (van Dijk & Kintsch, 1983). Still, most of the extensions to the original work generally expand the theory at the low-level. The two most significant high-level concepts which are still not handled are creative reading and treatment of the author's intent. It has been shown that creativity is linked to reading comprehension (Popov, 1993) and this has always been our contention. No other existing theory of reading has attempted to handle this aspect of reading, even though education has long recognized its significance. While the van Dijk and Kintsch theory may be valid for explaining structured reading tasks, it is unable to handle pleasure reading. Author intent is also rarely handled. Authors intend for text to be understood (for a recent discussion of this issue, see Gerrig, 1993). Therefore, there is a large amount of redundant information contained in text. The need to skim and skip portions of the text is imperative, but many other theories fail to allow for this. The coverage of reading theories we have handled is rather broad; we feel that many psychological theories and particularly artificial intelligence theories overlook the work done on reading in the field of education.

The second current influential model is the result of research done by Carpenter and Just (1988). They chose to concentrate on the working memory constraints of the reader and how these influence reading and recall issues. They, too, have had a great deal of success modeling human subjects' performance within their READER framework. The Carpenter and Just approach, however, also ignores issues of creative reading and author intent, as so many theories do. The primary focus of their work has always been on memory's influence on the reading process; in that regard, their work has had tremendous success and impact on future work. So, while the theory is capable of explaining much of what it set out to accomplish, it is insufficient for modeling the type of reading in which we are most interested.

The above whirlwind look at reading theories should have left the reader with a sense of discomfort. Each theory seems to miss something; curiously, what one theory is lacking seems to be taken up by another one. This is an important aspect of the past research to consider. Many of the theories above concentrate on the *what* of reading: what skills a reader needs and what types of comprehension exist. In contrast, the Kintsch and van Dijk model and the Just and Carpenter model both concentrate more on the *how* of reading. This dichotomy is partially the result of some of the theories being more heavily based in the education field, while others come from a background in psychology or cognitive science. While none of theories described were "wrong," none of them were complete enough. For one reason or another, the researchers chose to concentrate on a slice of the entire spectrum of what reading is. Some were concerned with what it is, some with how to teach it, some with why people read, and others with how the act is accomplished. A junior high school literature textbook (Beck et al., 1991) concisely sums up the above views. At the start of the text, the authors attempt to explain to the reader both how and why people read. They point out that "for many readers, the answer to that question [why people read] is, 'To meet new people. To have new experiences. To get information.'." This is a good summation of the *why*. The authors also attempt to touch on the how. They believe that if a reader is *actively* engaging the text, then they must be asking self-generated questions as they read the material, making predictions concerning what is going to happen next, and relating the text to their own experiences and life. Finally, when reading is finished, the active reader will attempt to put everything together, consider the hanging questions (if any), ponder on deeper meanings of the text, decide on a main idea, and so on. Our research attempts to explain how this is all accomplished.

10.3 A critique of past systems from the knowledge perspective

The knowledge grid representation can also be used to critically examine past systems. For example, Domeshek's ABBY system (1992) concentrated primarily on the social column of the grid. SAM and PAM (and most other early systems) were primarily interested in the physical and mental columns (with some items from the social added). CYRUS attempted to cover all aspects of the grid to at least a shallow level of understanding. Other attempts have tried to capture the essence of temporal relationships. This includes work in real temporal reasoning (see, for example, Scerba, 1979); for example, if you are told on Monday that something is due "next week," when is it really due? Before the Friday four days away? Within seven days? Within eleven days (the second Friday)? Social conventions and background influence this type of reasoning the most. Temporal research also included attempts to capture temporal metaphors, such as "March is rapidly approaching" (Lakoff & Johnson, 1980). Finally, there have been attempts to capture the information which would reside in the emotion column of our knowledge grid (e.g., O'Rorke & Ortony, 1992). The knowledge representation grid used in our theory synthesizes and extends these earlier attempts.

10.4 A critique of past systems from the creative understanding perspective

It is also possible to use the creative understanding perspective to evaluate past systems. While most systems restricted themselves to strict reliance on world knowledge, there were some exceptions. AQUA built explanations of scenarios which it did not possess knowledge of, such as terrorist suicide bombings. Meta-AQUA (Ram & Cox, 1993) attempts to modify its world knowledge in pursuit of the interpretation which holds more of the story coherently together. Looking over past systems, one can see that the ones which performed some level of creative understanding were presented as learning systems, for the most part. A person, it would seem, only learns when knowledge that they possess is challenged in some way by new material and an integration of the new material takes place.

11 Future work

On the side of implementation, current plans call for ISAAC to be extended and enhanced in a number of ways in the near future. First, the system will be expanded and refined to the point where a more detailed scenario understanding and more refined metacontrol are in place. These extensions will allow ISAAC to understand straightforward stories and will permit us to concentrate on the creative understanding aspects of our theory. The COMPERE–ISAAC connection will also be expanded to the point where COMPERE is fully integrated into the sentence processing supertask and all of the story is handled by COMPERE, with COMPERE receiving feedback from ISAAC concerning what it needs to perform next. Evaluation will begin on the system, as outlined above. We will also begin to extend ISAAC to work with additional stories, gradually building up the set of texts which it is capable of handling; for example, over the next few weeks, a second story will be brought into the realm of what ISAAC can read. Finally, on a more engineering note, ISAAC will be made easier to use and to understand through the addition of a graphical interface.

Within the framework of extending ISAAC as a reading system, we will also be extending its creative abilities. To increase the level of creativity, we will primarily be concentrating on the memory and explanation supertasks. The CUP algorithm will be fully implemented and tested in order to gain an understanding of where its power lies. A variety of tests will be performed with the creativity mechanisms at various levels performance. This will allow us to understand exactly what creativity gives ISAAC that was lacking in earlier systems.

Further work is also needed to refine both the reading theory in general and the creativity theory in particular. The increased level of implementation will allow this refinement to occur as we will be able to judge the success of ISAAC on more than a single story. It is also our intent to gather psychological data on aspects of the reading process and the creative understanding process which our theory predicts. This improvement in our theory, of course, will be channelled back into improved and more accurate implementation.

With the implementation increased to the above levels, we will be able to judge the validity of some of our theoretical claims. Various experimental tests will be performed on the system in order to gauge its abilities. *Ablation studies* (Cohen & Howe, 1988) will be done to pinpoint the impact each form of knowledge and each supertask has on the overall task of reading. These results will be compared to psychological data concerning the reading process. Finally, as mentioned earlier, another way which the ISAAC system will be judged is through the testing by high-school English teachers. Taken together, the metrics we have discussed will allow us to accurately and objectively judge ISAAC, as well as the theory of creative reading which drives it.

12 Conclusions

Reading is one of the most complex cognitive abilities we humans possess. Past research has generally concentrated on only a small portion of the entire process; this has led to a shallow understanding of the tasks involved. Our theory reflects this inherent complexity but better describes the comprehension process. By making heavy use of the total knowledge which exists within a story and by relying on an intensive interaction between the various reading tasks, our theory is capable of modeling the reading process. In addition to describing the reading task, we are also able to explain the inadequacies of previous systems.

The theory incorporates a general process of creative understanding as an integral part of reading. Creative understanding is modelled as an extension of mundane understanding. The power of the act comes from a willingness to not be bound by current world knowledge, to take a cognitive risk and attempt to understand something in a way it has never been understood before. This is accomplished through the techniques of imaginative memory, analogy, baseless analogy and problem reformulation. Function-driven morphological synthesis is the underlying mechanism which drives the creative understanding process. The knowledge organization scheme presented facilitates this process and acts as a rational bound on the degree of creative understanding which is attempted. Without the capability for creative understanding, a reading system will not be able to achieve the same level of reading that a human is capable of.

The task of reading is so complex and so interconnected, the idea of developing an understanding of the entire process seems, at times, a Sisyphean task. Perhaps this is why so many researchers are choosing to concentrate on smaller sections of the problem and are staying away from the global questions. That approach is respectable and will no doubt lead to significant results (as it already has in many cases). However, at some level of thought in the researcher's mind must exist a more global view of the process. Our theory provides such a view. We see it acting as a scaffold around which various other theories can be built and supported. We ourselves are working at a high point in the structure in the area of creative understanding.

The theory represents an integration and extension of a variety of results and evidence from past research and our own work. Through a combination of psychological insights and an analysis of past and current reading systems, we were able to develop the integrated, functional theory presented here. It is being implemented in the ISAAC system, where we intend to implement as much of the global theory as possible to more accurately model the human reading process, with emphasis placed on those aspects which will aid creative understanding.

13 Acknowledgements

This work is supported by a Fannie and John Hertz Foundation fellowship, by the National Science Foundation under grant IRI-8921256, by the Office of Naval Research under grant N00014-92-J-1234, and by the Georgia Institute of Technology. The authors would like to thank the members of IGOR (Anthony Francis, Mike Cox, Mark Devaney, and Paul Rowland), as well as Linda Wills, all of whom read numerous drafts and offered useful comments. An earlier form of this report appears in the PsychGrad Electronic Journal. The authors also thank the PsychGrad reviewers: Brian K. Smith, Paul Kershaw, and Christopher M. Hakala, as well as the editors: Zazie Todd and Matthew Simpson.

References

Allen, J. F. (1979). A plan-based approach to speech act recognition. Technical Report 131, University of Toronto.

Allen, J. F. (1983). Recognizing intentions from natural language utterances. In Brady, M. and Berwick, R. C., editors, *Computational Models of Discourse*. MIT Press, Cambridge.

Bailey, R. W. (1993). *Brothers of the Dragon*. Penguin Books, New York.

Beck, D. I. L., Lee, J. M., Brown, M., Goldenberg, P., Maley, A. P., Sohen, A. M., and Travers-Hovland, M. (1991). *Introduction to Literature*. Holt, Rinehart, and Winston, Inc., Orlando.

Birnbaum, L. (1986). *Integrated Processing in Planning and Understanding*. PhD thesis, Yale University. Research Report 489.

Black, J. B. and Seifert, C. M. (1981). The psychological study of story understanding. Technical Report 18, Yale University.

Bloch, A. (1963). Men Are Different. In Asimov, I. and Conklin, G., editors, *50 Short Science Fiction Tales*. MacMillan Publishing Co., New York.

Boden, M. A. (1991). *The Creative Mind: Myths and Mechanisms*. BasicBooks, New York.

Brown, F. (1992). Blood. In Asimov, I., Greensberg, M., and Olander, J. D., editors, *Microcosmic Tales: 100 Wondrous Science Fiction Short-Short Stories*. Daw Printing, New York.

Bulfinch, T. (1959). *Bulfinch's Mythology*. Dell Publishing, New York. Abridged edition by Edmund Fuller.

BusinessWeek (1993). Hot products: Smart design is the common thread. *BusinessWeek*, pages 54–78.

Carberry, M. S. (1988). Pragmatic modeling: Toward a robust natural language interface. *Computational Linguistics*, 3.

Carbonell, J. G. (1990). Derivational analogy: A theory of reconstructive problem solving and expertise acquisition. In Shavlik, J. W. and Dietterich, T. G., editors, *Readings in Machine Learning*. Morgan Kaufmann Publishers, Inc., San Mateo, CA.

Carpenter, P. A. and Just, M. A. (1988). The role of working memory in language comprehension. In Klahr, D. and Kotovsky, K., editors, *Complex Information Processing: The Impact of Herbert A. Simon*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Carpenter, T. and Alterman, R. (1993). Reading as a planned activity. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*, pages 300–305.

Clarke, A. C. (1962). *Profiles of the future*, chapter Hazards of Prophecy: The Failure of Imagination. Harper and Row, New York.

Clymer, T. (1968). What is reading? Some current concepts. In Robinson, H. M., editor, *Innovation and Change in Reading Instruction*. National Society for the Study of Education, Chicago.

Cohen, P. R. and Howe, A. E. (1988). How evaluation guides AI research. *AI Magazine*, 9(4):35–43.

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

Cullingford, R. (1981). SAM. In Schank, R. C. and Riesbeck, C. K., editors, *Inside Computer Understanding*, chapter 5. Lawrence Erlbaum Associates, Hillsdale, NJ.

Davis, F. B. (1968). Research in comprehension in reading. *Reading Research Quarterly*, 3:499–545. Summer.

DeJong, G. F. (1982). An overview of the FRUMP system. In Lehnert, W. G. and Ringle, M. H., editors, *Strategies for Natural Language Processing*, chapter 5, pages 149–276. Lawrence Erlbaum Associates, Hillsdale, NJ.

Domeshek, E. (1992). *Do the Right Thing: A Component Theory of Indexing Stories for Social Advice*. PhD thesis, Yale University, New Haven, CT.

Dyer, M. (1983). *In-Depth Understanding*. MIT Press, Cambridge.

Falkenhainer, B. (1987). Scientific theory formation through analogical reasoning. In *Proceedings of the Fourth International Workshop on Machine Learning*.

Fikes, R. E. and Nilsson, N. J. (1971). STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2:189–208.

Finke, R. A., Ward, T. B., and Smith, S. M. (1992). *Creative Cognition: Theory, Research, and Applications*. MIT Press, Cambridge.

Francis, A. G. (1993). Personal communication with Anthony Francis.

Gavelik, J. R. and Raphael, R. E. (1985). Metacognition, instruction and the role of questioning activities. In Forrest-Presley, D. L., MacKinnon, G. E., and Waller, T. G., editors, *Metacognition, Cognition, and Human Performance*, volume 2. Academic Press, New York.

Gentner, D. (1989). Mechanisms of analogical learning. In Vosniadou, S. and Ortony, A., editors, *Similarity and Analogical Reasoning*. Cambridge University Press, London.

Gerrig, R. J. (1993). *Experiencing Narrative Worlds: On the Pyschological Activities of Reading*. Yale University Press, New Haven.

Giuse, D. (1990). Efficient knowledge representation systems. *The Knowledge Engineering Review*, 5(1):35–50.

Goodman, M., Waterman, S., and Alterman, R. (1991). Interactive reasoning about spatial concepts. In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, pages 734–738.

Gough, P. B. (1972). One second of reading. In Kavanagh, J. F. and Mattingly, I. G., editors, *Language by Ear and by Eye*. MIT Press, Cambridge.

Graesser, A., Golding, J. M., and Long, D. L. (1991). Narrative representation and comprehension. In Barr, R., Kamil, M. L., Mosenthal, P., and Pearson, P. D., editors, *Handbook of reading research*, volume 2, chapter 8. Longman Publishing Group, White Plains, NY.

Graesser, A. C. (1981). *Prose Comprehension Beyond the Word*. Springer, New York.

Grice, H. P. (1957). Meaning. *Philosophical Review*, 67.

Grosz, B. J. and Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3).

Gruber, H. E. and Davis, S. E. (1988). Inching our way up mount olympus: The evolving-systems approach to creative thinking. In Sternberg, R. J., editor, *The Nature of Creativity*. Cambridge University Press, Cambridge.

Hammond, K. J. (1989). CHEF. In Reisbeck, C. and Schank, R., editors, *Inside Case-Based Reasoning*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Holbrook, J., Eiselt, K. P., and Mahesh, K. (1992). A unified model of syntactic and semantic error recovery in sentence processing. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, pages 195–200.

Huey, E. B. (1908). *The Psychology and Pedagogy of Reading*. The MacMillan Co., New York.

Jones, E. K. (1992). The flexible use of of abstract knowledge in planning. Technical Report 28, Northwestern University, The Institute for the Learning Sciences.

Just, M. A. and Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99(1):122–149.

Kass, A. M., Leake, D. B., and Owens, C. C. (1986). SWALE: A program that explains. In Schank, R. C., editor, *Explanation Patterns: Understanding Mechanically and Creatively*, pages 232–254. Lawrence Erlbaum Associates, Hillsdale, NJ. Appendix.

Kintsch, W. and van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5):363–394.

Kolodner, J. (1984). *Retrieval and Organization Strategies in Conceptual Memory: A Computer Model*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Kolodner, J. L. (1993). *Case-Based Reasoning*. Morgan-Kaufman Publishers, Inc., San Mateo, CA.

Kolodner, J. L. and Wills, L. M. (1993). Paying attention to the right thing: Issues of focus in case-based creative design. In *Proceedings AAAI-93 CBR Workshop*.

Lakoff, G. and Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press, Chicago, IL.

Lebowitz, M. (1983). Memory-based parsing. *Artificial Intelligence*, 21.

Lehnert, W. (1977). *The Process of Question Answering*. PhD thesis, Yale University. Research Report 88.

Lenat, D. (1990). The ubiquity of discovery. In Shavlik, J. W. and Dietterich, T. G., editors, *Readings in Machine Learning*. Morgan Kaufmann Publishers, Inc., San Mateo, CA.

Levy, D. M. (1979). Communicative goals and strategies: Between discourse and syntax. In Givon, T., editor, *Discourse and Syntax*. Academic Press, New York.

Mackworth, J. F. (1972). Some models of the reading process: Learners and skilled readers. *Reading Research*

Quarterly, 7:701–733.

Mahesh, K. (1993). A theory of interaction and independence in sentence understanding. Technical Report GIT-CC-93/34, Georgia Institute of Technology.

Newell, A. and Simon, H. A. (1972). *Human Problem Solving*. Prentice-Hall, Englewood Cliffs, NJ.

Nisbett, R. E. and Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3).

Niven, L. (1973). *The Flight of the Horse*. Ballantine Books, New York.

O'Rorke, P. and Ortony, A. (1992). Abductive explanation of emotions. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, pages 594–598.

Orwell, G. (1949). *1984*. Harcourt Brace Jovanovich, Inc., New York.

Popov, A. (1993). Creativity and reading comprehension. *Journal of Creative Behavior*, 26(3):206–212.

Pratchett, T. (1983). *The Colour of Magic*. Penguin Books, New York.

Ram, A. (1991). A theory of questions and question asking. *The Journal of the Learning Sciences*, 1(3&4):273–318.

Ram, A. and Cox, M. (1993). Introspective reasoning using meta-explanations for multistrategy learning. In Michalski, R. and Tecuci, G., editors, *Machine Learning: A Multistrategy Approach*, volume IV. Morgan Kaufman Publishers, San Mateo, CA. In Press.

Ram, A. and Hunter, L. (1992). The use of explicit goals for knowledge to guide inference and learning. *Applied Intelligence*, 2(1):47–73.

Robeck, M. C. and Wilson, J. A. R. (1974). *Psychology of Reading: Foundations of Instructions*. John Wiley and Sons, Inc., New York. BF 456 .R2.

Rumelhart, D. E. (1977). Understanding and summarizing brief stories. In Berge, D. L. and Samuels, J., editors, *Basic processes in reading and comprehension*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Sacerdoti, E. D. (1974). Planning in a hierarchy of abstraction spaces. *Artificial Intelligence*, 5(2):115–135.

Sacerdoti, E. D. (1977). *A Structure for Plans and Behavior*. American Elsevier, New York.

Scerba, T. M. (1979). Human and computational models of time cognition. Masters thesis, Yale University.

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

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

Schank, R., Goldman, N., Rieger, C., and Riesbeck, C. (1975). Inference and paraphrase by computer. *Journal of the Association for Computing Machinery*, 22(3):309–328.

Schneider, W. (1985). Developmental trends in the metamemory-memory behavior relationship: An integrative review. In Forrest-Presley, D. L., MacKinnon, G. E., and Waller, T. G., editors, *Metacognition, Cognition, and Human Performance*, volume 1. Academic Press, New York.

Sidner, C. L. (1985). Plan parsing for intended response recognition in discourse. *Computational Intelligence*, 1(1).

Smith, N. B. (1986). *American Reading Instruction*. International Reading Association, Newark, Delaware.

Strang, R., McCullough, C. M., and Traxler, A. E. (1967). *The Improvement of Reading*. McGraw-Hill Book Company, New York, fourth edition.

Sussman, G. J. (1973). A computational model of skill acquisition. Technical Report AI TR-297, MIT.

Taylor, C. W. (1988). Various approaches to and definitions of creativity. In Sternberg, R. J., editor, *The Nature of Creativity*. Cambridge University Press, Cambridge.

Turner, S. (1992). *MINSTREL: A Computer Model of Creativity and Storytelling*. PhD thesis, University of California, Los Angeles.

van Dijk, T. A. and Kintsch, W. (1983). *Strategies of Discourse Comprehension*. Academic Press, New York.

Wallas, G. (1926). *The Art of Thought*. Harcourt Brace Jovanovich, Inc., New York.

Weinert, F. E. (1987). Introduction and overview: Metacognition and motivation as determinants of effective learning and understanding. In Weinert, F. E. and Kluwe, R. H., editors, *Metacognition, Motivation, and Understanding*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Weisberg, R. W. (1986). *Creativity. Genius and Other Myths*. Freeman, New York.

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, pages 94–104.

Wellman, H. M. (1985). The origins of metacognition. In Forrest-Presley, D. L., MacKinnon, G. E., and Waller, T. G., editors, *Metacognition, Cognition, and Human Performance*, volume 1. Academic Press, New York.

Wilensky, R. (1981). PAM. In Schank, R. C. and Riesbeck, C. K., editors, *Inside Computer Understanding*, chapter 7. Lawrence Erlbaum Associates, Hillsdale, NJ.

Wilensky, R. (1983). *Planning and Understanding*. Addison-Wesley, Reading, MA.

WKRP Episode 60 (1980). Venus and The Man. Written by Hugh Wilson, Directed by Rod Daniel.

Woods, W. A. (1986). Transition network grammars for natural language analysis. In Grosz, B. J., Jones, K. S., and Webber, B. L., editors, *Readings in Natural Language Processing*. Morgan Kaufmann Publishers, Inc, Los Altos.