31st Annual Association for Computing Machinery Southeast Conference, Birmingham, AL, April 1993.

A New Perspective on Story Understanding *

Kenneth Moorman and Ashwin Ram
{kennethm,ashwin}@cc.gatech.edu
(404) 853-9381
Georgia Institute of Technology
College of Computing
Atlanta, GA 30332-0280

Abstract

Reading has been studied for decades by both artificial intelligent researchers and by cognitive psychologists. Yet there still does not exist a theory which accurately describes the process nor a computer system which performs it. To overcome the shortfalls of previous systems, we have developed a functional theory of the reading comprehension process. The theory not only enables us to implement a model of the reading task, the ISAAC (Integrated Story Analysis And Comprehension) system; it also allows us to analyze past systems and their problems more completely and in a common framework. We also use this paper to present our general research philosophy, the current level of the implementation of our ideas, and our plans for future research.

1 Introduction

No one would dispute the claim that a majority of people read a great deal over the course of a day. In fact, reading seems to be a hallmark of intelligent behavior. Perhaps this is why so much interest has been given to the process, both by psychologists and by researchers in artificial intelligence. Psychologists have tried to develop a comprehensive theory of reading, while artificial intelligence researchers have attempted to create a reading system. These earlier theories and models failed for a variety of diverse reasons which are tied to a central point—the models in question did not make use of a complete theory of reading. As a result, the various attempts only concentrated on a portion of the problem. For example, there were systems which were syntax-driven, such as Woods' ATN systems [11]; these did not understand text at a high level of comprehension. Systems which did strive for comprehension level often sacrificed range, such as BORIS [5], which understood two researcher-created stories in-depth.

We have developed, and are continuing to refine, a theory of story understanding which enables us to overcome some of the problems of these earlier systems, We are approaching the problem from an integrated perspective, which will enable our system to succeed in ways which previous ones failed. The ISAAC (Integrated Story Analysis And Comprehension) system will be capable of reading a variety of texts in their unedited forms. Additionally, ISAAC will be able to understand the stories that it reads at a level not possible before. As part of our theory, we have developed the concept of *creative understanding*. This is one reason we selected science fiction stories as ISAAC's primary reading domain, as these offer unique opportunities for this type of comprehension.

2 Our approach

After a review of previous story understanding programs, we realized that the systems were unclear concerning what they were trying to accomplish. While presented as story understanding systems, they actually appeared more like "episode understanding" systems. These programs did not use the form of the story as an aid to understanding; instead, the text was simply seen as the written description of a real-world event. It was this real-world event, or *episode*, that the system was to understand.

As a result, we began to examine more carefully the concept of story understanding. This examination of the problem led us to the following central idea: 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 a different aspect of the global reading process.

To date, we have identified six supertasks necessary for textual understanding. Each supertask is made up of numerous tasks which tightly interact within the function of the supertask. These supertasks are:

- **metacontrol**: Metacontrol integrates the other supertasks. *Focus control* manages the depth of reading

*This work was supported by a Fannie and John Hertz Foundation fellowship and by the Georgia Institute of Technology

based on interest and understanding. *Time management* allows the reader to make decisions based on time resources. Another task is *suspension of disbelief* which enables the reader to accept, at least temporarily, a text which violates her/his world view.

- **language understanding**: This supertask is responsible for "low level" understanding, such as *pronoun reference*, *syntactic parsing*, and *lexical retrieval*. Other tasks are *punctuation analysis*, which reasons about punctuation of phrases; and *tense analysis* which attempts to discover the text's tense.
- **story structure understanding**: This handles the details of the text which relate to the story structure, such as *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 identifies the category of text the story belongs to.
- **episodic understanding**: The tasks making up episodic 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 which the agents are involved with.
- **explanation and reasoning**: This supertask performs high-level reasoning and learning. *Creative analogy* 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 meanings to words not in the lexicon. *Interest management* controls the reader's level of interest in the story. *Belief management* manages the beliefs of the agents involved in the episode. 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 construction*, which creates the various cases which result from a reading experience; *memory retrieval*, which returns information from memory; and *memory storage*, which places new information and cases into memory.

3 Benefits of our methodology

What does our approach provide that the older approaches did not? First, each of the supertasks is customizable through the use of *models* of what to expect from a given genre of text. For example, if ISAAC is aware that the text that it is reading is a newspaper article, it will attempt to read it differently than if it realizes it is reading a fictional story. This flexibility will allow ISAAC to handle numerous text styles. A second benefit of our modular approach is the potential expansion of certain components beyond the level of text understanding. What we call the *episodic understander* could be a general purpose module which could someday interpret what it "sees" going on around it, in addition to what it sees in a piece of text. Next, the metacontrol-based integration of the various components of the ISAAC system allows the system to guide its own reading processes. Rather than attempt to read everything at the most intensive depth of understanding, ISAAC will alter its reading depth based on the current needs of the active tasks. This will enable ISAAC to skim sections which are uninteresting or which contain known information, and will allow it to read longer passages than possible with earlier systems.

Finally, our theory enables us to efficiently examine older reading systems and see why they were unable to handle the complete reading task. Early systems such as SAM [4] and PAM [10] were attempts to handle episodic understanding through scripts and plans. Dyer's BORIS system was an integrated approach which tried to bring together the current theories of its time; unfortunately, it became too much of an engineering task to give BORIS the knowledge it needed for more than two short stories. CYRUS [6] was an implementation of the memory management supertask but sacrificed the rest of the reading process. The later system, AQUA [7], contributed to the explanation and reasoning supertask and somewhat to the memory management supertask, but it ignored certain story structure aspects which could have aided the comprehension process. Other systems concentrated on syntactic issues, such as ATNs, which are part of our language understanding supertask; these were generally too low-level to get much past the level of sentence comprehension. Finally, there was one field of research which focused on story structure as the key to understanding, e.g. story grammars [8]. High-level reasoning was where these systems fell short, relying too much on the story structure alone.

These past systems all ignored various aspects of the total process of reading, leading to them having to work harder at the portions on which they did concentrate. Our theory has tried to address the deficiencies of past systems by incorporating the knowledge gained from each attempt. The theory is not simply a strict union of these past theories, however. Our integration allows each supertask to support the others, causing the whole to be more powerful than any single component.

4 Story understanding and creativity

During our research, we have noticed an interesting correlation between people who are creative and people who are widely read. This led us to consider if ISAAC could possess artificial creativity. One thing that creative people are able to do is apply past experiences in new and different ways to current problems. This application often leads to a creative solution being generated. ISAAC will have access to past reading experiences and be able to reason about them in much the same way. These creative applica-

	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 1: Knowledge representation grid

tions of past experience will happen in two primary fashions. First, ISAAC will be able to directly adapt solutions from past episodes to fit the current problem. Second, ISAAC will be able to use past creative experiences in much the same fashion as Carbonell's *derivational replay* [2] in order to produce creative solutions. Both of these techniques will enable ISAAC to act in a creative fashion during the reading experience, allowing it to come to a *creative understanding* of the story.

On the other hand, it is not enough simply to allow a reasoning system to make any adaptations that it chooses to. This could lead to totally nonsensical concepts being generated. However, one does not wish to discount the possibility of some type of modification from existing; to do this would be to constrain the potential creativity of the system. To satisfy these conflicting goals, we employ a multidimensional grid onto which all knowledge is mapped (see Figure 1). One dimension is a traditional, Schankian breakdown of knowledge [9] into *action*, *object*, *agent*, and *state*. The other is a natural breakdown according to mode of existence—*physical*, *mental*, *emotional*, *social*, and *temporal*. Concepts are located within a cell of the grid according to their descriptions. For example, walking is a physical action while hating is an emotional one. Concepts are not restricted to one cell within the grid; a person, for example, is a physical agent in some cases, a physical object in others, and a mental agents in the mind of another agent. The constraints on the adjustments comes from the fact that movement within the grid is a costly process. A movement of a concept within a grid cell is the cheapest type of transformation to make, movement along either a row or a column is more expensive, and movement along both axes is the most expensive type of movement of all.

In addition to using past reading episode to aid creativity, creativity is itself extremely useful in the process of story understanding. Consider the problem of reading a story about time travel. The part of the reader's mind which is anchored in reality understands that time travel is impossible at this point in time. Yet, to make sense of the story, it is necessary to make use of a *willing suspension of disbelief* [3]. To some degree, this must occur in all reading. No one really believes that the defense attorney's investigator is able to find the crucial piece of data just in time to save the client's case; it just makes the story a nicer one to read. This suspension of disbelief is strongest in the science fiction domain, which is one reason we chose it for our research.

5 An example

The following is an excerpt from a short story we are working with, *Men Are Different* by Alan Bloch [1].

> 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.
>
> 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 that I met the Man. He must have been the last Man in this system. 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.

5.1 Discussion

Consider the understanding of this story from the perspective of our theory. The *pronoun reference* task determines that the *I* in the story refers to an agent whose occupation is archaeology. *Punctuation analysis* "notices" that *Men* is capitalized in an unusual way, implying that the word carries some special meaning in the story's context. *Character identification* contributes the knowledge that *I* is a story character. This enables the *genre identifier* to guess that the story is a first-person narrative, since *I* is acting as the subject of a sentence. By the end of the first paragraph, the story structure model knows that the main character of this first-person narrative is a robot archaeologist. and the episode model knows that the robot is wondering about the fate of mankind. These models continue to be refined over the course of the story.

During the reading experience, multiple cases are being built. First, there is the case representing the events of

the story, which is a summary of what the reader understands of the text. Second, this case contains a subcase representing the flashback scene at the end of the tale. Third, there is a case which represents the actions of the reader as it read and understood the story; this case would contain, for example, information concerning when the reader realized that the narrator was a robot. It is a trace of what the reader went through during comprehension rather than simply being the result of the comprehension task. If the reader greatly empathized with the robotic narrator, a fourth case may also exist. This would be a "pseudo-experience" in the memory of the reader—s/he would remember it as being in the context of a story, but the episode could aid the reader if s/he ever has to deal with a person from a radically different culture and if s/he remembers that one should not project one's way of life onto another due to surface similarities.

Finally, if the reader is familiar with science fiction stories, the start of this reading experience will trigger remindings which aid the process. If the reader is comfortable with science fiction stories, the idea of a robot narrator will not require as much disbelief suspension as needed by a science fiction novice. For a reader to comprehend the story fully, s/he 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. If the reader refuses to accept this story world, the story will not be fully appreciated.

6 Current and future work

ISAAC is currently implemented at a very basic level of functionality, with the intent that the system be used as a testbed for our ideas. The system contains enough information to build a coherent story model of the short science fiction story discussed above. ISAAC also performs rudimentary pronoun referencing and handles the creation of a limited episodic model of the events depicted in the story. Finally, it maintains a reasoning trace of its own activities. While the implementation is still at a beginning level, what we have accomplished so far is important for several reasons. It has helped to show that the approach we advocate is at least as competent as older theories. It has allowed us to gain a concrete understanding of what each of the supertasks in a reading system needs to be. Finally, it has permitted us to glimpse the complex interactions which exist between the modules of the reading task. The short-term goal is for ISAAC to be able to read a few text passages from several genres. After the general background knowledge is in place, expanding ISAAC to a new genre should be relatively straightforward, due to its modular nature. The ultimate goal is for ISAAC to be able to read stories from several diverse genres, perform in-depth reasoning, perform creative reasoning, and do all of this in a dynamic, reader-driven fashion.

7 Conclusion

ISAAC represents the next generation of story understanding systems. It does not separate the reading process from the rest of cognition; instead, ISAAC sees reading as one aspect of a general cognitive model. The reading it performs will be at a better level and scaling to larger reading tasks should be easier than with older systems. The power of ISAAC comes from our theory of what the reading process entails. How do people read? We feel that the answer to this question will lead AI a step closer to the ultimate goal of an intelligent computer system.

Beyond that, our methodology allows us to critically examine past story understanding theories and systems within a common framework and to see where they fell short. Out hope is that the awareness we possess of the total comprehension process will enable us to not make some of the sacrifices earlier systems were forced to.

References

[1] BLOCH, A. Men Are Different. In *50 Short Science Fiction Tales*, I. Asimov and G. Conklin, Eds. Macmillan Publishing Company, New York, 1963.

[2] CARBONELL, J. G. Derivational analogy: A theory of reconstructive problem solving and expertise acquisition. In *Readings in Machine Learning*, J. W. Shavlik and T. G. Dietterich, Eds. Morgan Kaufmann Publishers, Inc., San Mateo, California, 1990.

[3] CORRIGAN, R. W. *The World of the Theatre*. Scott, Foresman and Company, Glenview, IL, 1979.

[4] CULLINGFORD, R. SAM. In *Inside Computer Understanding*, R. C. Schank and C. K. Riesbeck, Eds. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey, 1981, ch. 5.

[5] DYER, M. *In-Depth Understanding*. MIT Press, Cambridge, 1983.

[6] KOLODNER, J. *Retrieval and Organization Strategies in Conceptual Memory: A Computer Model*. Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey, 1984.

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

[8] RUMELHART, D. E. Understanding and Summarizing Brief Stories. In *Basic Processes in reading and comprehension*, D. L. Berge and J. Samuels, Eds. Lawrence Erlbaum Associates, Hillsdale, NJ, 1977.

[9] SCHANK, R., AND ABELSON, R. *Scripts. Plans, Goals, and Understanding*. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey, 1977.

[10] WILENSKY, R. PAM. In *Inside Computer Understanding*, R. C. Schank and C. K. Riesbeck, Eds. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey, 1981, ch. 7.

[11] WOODS, W. A. Transition Network Grammars for Natural Language Analysis. In *Readings in Natural Language Processing*, B. J. Grosz, K. S. Jones, and B. L. Webber, Eds. Morgan Kaufmann Publishers, Inc, Los Altos, 1986.