Recovering from Erroneous Inferences

Kurt P. Eiselt*
Department of Information and Computer Science
University of California
Irvine, California 92717

Abstract

Many models of natural language understanding make inference decisions as they process a text, but few models can correct their interpretation when later text reveals that earlier inference decisions are wrong. This paper describes how ATLAST, a marker-passing model of text understanding, addresses this problem. The keys to ATLAST's error recovery capability are a means for remembering the choices it could have made but did not, and a means for initiating the re-evaluation of those previously rejected choices at the appropriate times. This paper also discusses some of the arguments for and against the psychological validity of a theory of inference retention in human text understanding.

I. Introduction

As we read, we make unconscious decisions about the meaning of ambiguous words, sentences, or passages based on incomplete information. Often those decisions are wrong and we must revise our understanding of the text. For example, consider the following simple story:

```
Text 1: Fred asked Wilma to marry him.
        Wilma began to cry.
```

Interpreting this text requires that a causal relationship between Fred's proposal and Wilma's tears be inferred. One such possible relationship is that Wilma was happy about Fred's proposal and was crying "tears of joy." Another equally likely inference is that Wilma was crying because she was saddened or upset by the proposal.[1]

Now consider this variation of Text 1:

```
Text 2: Fred asked Wilma to marry him.
        Wilma began to cry.
        She was saddened by the proposal.
```

Assuming that after processing the first two sentences of Text 2, the text understander has inferred that Wilma is happy, how does the understander resolve that inference with the contradictory third sentence?

One solution is to postpone making inferences for as long as possible so that potential conflicts are resolved before any decisions are made. However, this solution becomes less viable as texts increase in length. A better solution is to make inferences as the opportunities arise, then revise initial inferences if later text shows them to be incorrect. This paper describes how one model of text understanding, ATLAST, simplifies the error recovery process by remembering the alternative inferences it could have made but did not, and reconsidering those alternatives when subsequent text suggests they might now be correct.

Most models of language understanding fail to address the problem of recovery from erroneous inferences, but there have been exceptions. Granger's ARTHUR (1980) was able to supplant incorrect inferences by maintaining a map of pointers to all inferences generated during the processing of a text, whether or not they appeared in the final representation. O'Rorke (1983) designed a story understander called RESUND that used non-monotonic dependencies to correct false assumptions. Norvig's FAUSTUS (1983) temporarily stored rejected inferences using a process similar to the retention process discussed in this paper. FAUSTUS represented inferences as frames, and rejected frames were stored in a separate data base in case later text forced revision of earlier decisions.

II. How it Works

ATLAST's ability to revise its interpretation of a text depends in large part on the use of a relational network to represent knowledge. ATLAST uses marker-passing to search its relational network for paths that connect meanings of open-class words from the input text. A single path is a chain of nodes, representing objects or events, connected by links, corresponding to relationships between the nodes. Any nodes in a path which are not explicitly mentioned in the text are events or objects that are inferred; therefore, these paths are called *inference paths.* A set of inference paths that joins all words in the text into a connected graph represents one possible interpretation of the text. In this respect ATLAST resembles a number of other models of text understanding that utilize marker-passing or spreading activation (e.g., Charniak, 1983; Cottrell, 1984; Hirst,

*This research was supported in part by the National Science Foundation under grants IST-81-20685 and IST-85-12419 and by the Naval Ocean Systems Center under contracts N00123-81-C-1078 and N66001-83-C-0255.

[1]Experimental evidence indicates that either interpretation is equally likely when this text is presented to human subjects (Granger & Holbrook, 1983).

1984; Quillian, 1969; Riesbeck & Martin, 1986; Waltz & Pollack, 1985). The paths that make up the current interpretation are called *active paths.*

For any given text, however, there may be a great number of possible interpretations, many of which are nonsensical. The problem then is determining which of the possible interpretations provides the best explanation of the text. ATLAST deals with this problem by applying inference evaluation metrics. These metrics are used to compare two competing inference paths and select the more appropriate one. Two inference paths compete when they connect the same two nodes in the relational network via different combinations of links and nodes. The path that fits better with the current interpretation is activated (i.e., it becomes part of the interpretation). The other path is de-activated but not discarded. Instead, that path is *retained* in order to facilitate error recovery as described below. The choice of one inference path over another is made as soon as ATLAST discovers that the two paths compete; ATLAST does not postpone inference decisions. As the marker-passing search mechanism finds more paths, ATLAST constructs an interpretation consisting of those paths which survive the evaluation process. When the marker-passing and evaluation processes end, the surviving active paths make up the final interpretation of the text.

In addition to the assumption of a specific representation scheme, ATLAST relies on two key processing features for error recovery: the ability to remember inference paths that it originally decided should not be part of its interpretation of the input text, and a mechanism for recognizing when these rejected paths should be reconsidered. Without a mechanism for knowing when and how to re-evaluate the retained paths, the retention feature alone provides no benefit.

There are two ways in which the re-evaluation of a retained path can be initiated. The first is through direct rediscovery of the retained path by the search process. Because the passing of markers begins in different places at different times during the processing of text, the same inference path may be discovered (or more appropriately, rediscovered) more than once. If a rediscovered path is not currently part of ATLAST's interpretation of the text (i.e., the path has been discovered earlier, rejected by the evaluation metrics, but retained), that path is re-evaluated against the competing path which is part of the interpretation. This rediscovery process initiates reconsideration of some of the retained paths, but it is not dependent upon retention because these paths would be reconsidered even if they had not been retained.

Some retained paths, though, will not be rediscovered, but the inferences made from later text may change the interpretation in such a way that these paths now should be included. ATLAST uses a method of "piggy-backing" the re-evaluation of these paths onto the evaluation of paths which are directly discovered or rediscovered by the search process. If a (re)discovered path is evaluated against a competing path in the current interpretation, any subpaths or superpaths of the (re)discovered path are also evaluated against the current interpretation. In this way, ATLAST attempts to limit re-evaluation to those paths that are currently relevant.[2] Without the ability to force re-evaluation of paths rejected early in processing but not rediscovered later, ATLAST's final interpretation probably will be incorrect. Indirectly initiating the re-evaluation of previously rejected inference paths is essential to ATLAST's error recovery capability and is dependent upon inference retention.

III. ATLAST in Action

An example of ATLAST processing a simple but potentially misleading text will illustrate the program's capacity for error recovery. This section describes the operation of ATLAST as it arrives at an interpretation for a simplified version of Text 2:

```
Text 3: Fred proposed.
        Wilma cried.
        Wilma was sad.
```

Although this is a simplified version of the original text (because ATLAST's syntactic abilities are limited), the relevant inference decisions should be the same for both texts. In the following example, many of the steps are left out for the sake of brevity. The corresponding memory structure is shown in Figure 1.

As ATLAST reads the first sentence from left to right, it finds a path from `proposed` to `Fred`. At this point, there is no candidate interpretation for the text, thus no competing inference paths, so this path becomes the first member of the set of active paths:

```
path0 from PROPOSE-MARRIAGE to FRED:
  PROPOSE-MARRIAGE has the role-filler
    GENERIC-HUMAN
  GENERIC-HUMAN has the instance FRED
  activating path0
```

While processing the second sentence of the text, ATLAST finds a path denoting a causal relationship between `proposed` and `cried`. This path represents the inference that the crying results from a state of happiness which in turn results from the proposal of marriage. This path is added to the set of active paths:

[2] Constraining reconsideration to just those paths that completely contain or are completely contained by the (re)discovered path has proven to be too restrictive for another sample text. In that case, one retained path which should have been part of the final representation was neither directly nor indirectly chosen for re-evaluation. Relaxing the constraints allowed ATLAST to recover while still avoiding the re-evaluation of every retained path.

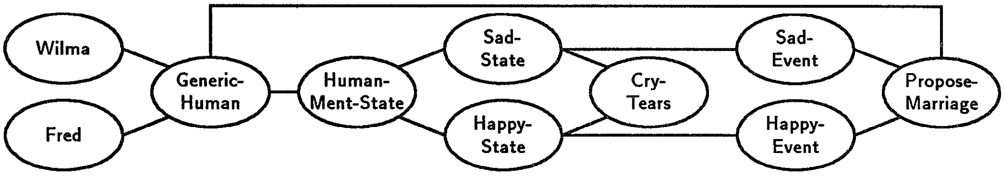

Figure 1: The organization of nodes in the memory structure for Text 3.

```
path4 from CRY-TEARS to PROPOSE-MARRIAGE:
  CRY-TEARS is a result of HAPPY-STATE
  HAPPY-STATE is a result of HAPPY-EVENT
  HAPPY-EVENT has the instance PROPOSE-MARRIAGE
  activating path4
```

Next, ATLAST discovers a path that provides an alternate interpretation to that offered by the previous path. During this example, ATLAST was instructed to give preference to older paths over newer paths when no other evaluation metric was able to make a decision.[3] Thus, the newer path is not added to the set of active paths:

```
path5 from CRY-TEARS to PROPOSE-MARRIAGE:
  CRY-TEARS is a result of SAD-STATE
  SAD-STATE is a result of SAD-EVENT
  SAD-EVENT has the instance PROPOSE-MARRIAGE
path4 older than path5
  de-activating path5
```

ATLAST now finds a path that connects `cried` to `Wilma` and adds it to the set of active paths:

```
path9 from CRY-TEARS to WILMA:
  CRY-TEARS is a result of SAD-STATE
  SAD-STATE is an instance of HUMAN-MENT-STATE
  HUMAN-MENT-STATE is an attribute of
    GENERIC-HUMAN
  GENERIC-HUMAN has the instance WILMA
  activating path9
```

The interpretation now contains three paths: path 0, path 4, and path 9. There is a semantic contradiction among the active paths at this time in that path 9 is an inference that Wilma cried because she was sad while path 4 says that the tears were shed due to a state of happiness induced by the marriage proposal. ATLAST does not notice the contradiction because the two paths are not competing paths. This is the best interpretation based on the paths discovered so far. ATLAST then finds a competing path from `Wilma` to `cried`. This new path, path 11, shares more nodes with other active paths than does its competing path, path 9; this is one of the criteria employed to decide which path explains more of the input. In this case, path 11 explains more input so it is added to the set of active paths and path 9 is moved to the set of retained paths:

```
path11 from CRY-TEARS to WILMA:
  CRY-TEARS is a result of HAPPY-STATE
  HAPPY-STATE is an instance of HUMAN-MENT-STATE
  HUMAN-MENT-STATE is an attribute of
    GENERIC-HUMAN
  GENERIC-HUMAN has the instance WILMA
path11 has more shared nodes than path9
  de-activating path9
  activating path11
```

As the final sentence is processed, ATLAST discovers a path connecting `proposed` to `sad`. This path is added to the set of active paths. In addition, this new path has four superpaths among the set of retained paths, and these paths are re-evaluated. One of these superpaths, path 5, is now preferred over the active path 4 because it is reinforced by path 15 (i.e., it contains the active path 15 as a subpath). Path 4 is moved from the set of active paths to the retained paths, and path 5 is moved from the retained paths to the active paths:

```
path15 from SAD-STATE to PROPOSE-MARRIAGE:
  SAD-STATE is a result of SAD-EVENT
  SAD-EVENT has the instance PROPOSE-MARRIAGE
also reconsidering: (path13 path10 path5 path2)
  activating path15
path11 shorter than path13
  de-activating path13
path4 shorter than path10
  de-activating path10
path5 has more shared nodes than path4
  de-activating path4
  activating path5
path0 shorter than path2
  de-activating path2
```

[3]This tendency to prefer older inferences over newer ones results from the work on differences in human inference decision behavior noted in an earlier footnote. The theory that was proposed to explain the differences suggests that some subjects prefer older inferences when faced with a choice between competing inferences, while other subjects prefer newer inferences. The people who prefer older inferences are called "perseverers" while those who prefer newer inferences are called "recencies." ATLAST is capable of modeling either kind of behavior by changing one of its evaluation metrics; it recovers from erroneous inferences in either mode.

The previous step demonstrates the need for inference path retention. Path 5 has been found directly several times prior to this point. Each time, the evaluation metrics have determined that path 4 fits better with the context. Now that path 15 is part of that context, path 5 is determined to be more appropriate than path 4. Had path 5 not been retained after being rejected earlier, it could not have been reconsidered at this time, nor would it ever have been reconsidered because the search process will not find path 5 again. If path 5 had not been retained, path 4 would incorrectly end up in the final representation of the story. In fact, this is what happens when ATLAST's retention capability is disabled while processing Text 3.

Continuing with the example, ATLAST finds a new path from `cried` to `sad` and adds it to the active paths. This new path also forces the reconsideration of several retained superpaths, including path 9, which is now preferred over its old competitor, path 11, because path 9 now shares more nodes with other active paths than does path 11. Path 9 is returned to the set of active paths and path 11 becomes a retained path, again illustrating the usefulness of inference retention:

```
path18 from SAD-STATE to CRY-TEARS:
  SAD-STATE has the result CRY-TEARS
also reconsidering: (path16 path13 path9 path14
  path7 path6 path3)
  activating path18
path15 shorter than path16
  de-activating path16
path11 shorter than path13
  de-activating path13
path9 has more shared nodes than path11
  de-activating path11
  activating path9
          .
          .
          .
```

ATLAST then discovers the last new path to be added to the set of active paths. This path connects `Wilma` and `sad`.

```
path20 from SAD-STATE to WILMA:
  SAD-STATE is an instance of HUMAN-MENT-STATE
  HUMAN-MENT-STATE is an attribute of
    GENERIC-HUMAN
  GENERIC-HUMAN has the instance WILMA
  activating path20
```

The marker-passing mechanism will uncover nine more new paths to be considered and rediscover many others; these paths will in turn force the re-evaluation of a number of retained subpaths and superpaths of those paths. However, none of these paths will be incorporated into the final interpretation of the text, which consists of paths 0, 5, 9, 15, 18, and 20.

IV. Retention Issues

The principle of retaining rejected inference paths is inspired by experimental work which has led to a theory of lexical disambiguation called *conditional retention* (Granger, Holbrook, & Eiselt, 1984). According to this theory, lexical disambiguation is an automatic process in which all meanings of an ambiguous word are retrieved, the meaning most appropriate to the preceding context is chosen, and the other meanings are temporarily retained. In the case where the ambiguous word appears within a short text, the meanings are retained until the end of the text. Should later text contradict the initially chosen meaning, the retained meanings for that word are reconsidered in light of the updated context, and a new meaning is selected without repeating the lexical retrieval process. The theory of conditional retention thus offers an explanation of how readers can recover from an incorrect choice of word meaning without reprocessing the text. Because the choice of a word meaning will affect the inferences which are made during the understanding of a text, the theory of conditional retention has implications for making inference decisions at levels other than the lexical level. Following this assumption, ATLAST uses the inference retention mechanism described in Sections II and III to recover from both incorrect lexical inferences as well as erroneous pragmatic inferences.

However, the theory of conditional retention is by no means widely accepted, and the criticisms of conditional retention should be taken into consideration when evaluating ATLAST's utility as a cognitive model. One argument against conditional retention is a large body of experimental evidence which shows that, almost immediately after a meaning of an ambiguous word has been selected, the alternate meanings seem as if they had never been recalled (e.g., Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982). This has been interpreted by some as proof that retention does not occur. On the other hand, these experiments were not specifically designed to look for evidence of retention. Also, as shown by Holbrook, Eiselt, Granger, and Matthei (1987), the results of some experiments (e.g., Hudson & Tanenhaus, 1984) can be interpreted in such a way as to support the theory of conditional retention, though not conclusively. The one experiment to date that was designed to look for retention (Granger et al., 1984) also yielded inconclusive results.

A frequent and deserved criticism of the conditional retention theory is that it offers no concrete answer to the question of how long alternate choices are retained; it says only that the choices are retained until the end of the text if the text is short. The experiment described by Granger et al. (1984) did not address this issue, but new work with ATLAST may suggest some answers. ATLAST has been modified so that a path is given a time stamp indicating the time at which it was added to the set of retained paths. In addition, a limit has been placed on the amount of time that a path can be retained without being reconsidered. With these modifications, the minimum duration of retention that is sufficient to allow ATLAST to arrive at the correct interpretation of a given text can be determined empirically. This in turn will enable us to investigate, for example, the possibility of a correlation between the duration of retention and structural cues such as clause boundaries. If interesting predictions do arise from this work, it may be possible to test these predictions in the laboratory with human subjects.

Another problem with the conditional retention theory is that it assumes human readers recover from errors without rereading the text. However, as Carpenter and Daneman (1981) demonstrate through studies of eye fixations of human subjects while reading, there are texts that cause a reader to backtrack when a semantic inconsistency is discovered in an ambiguous text. Carpenter and Daneman propose that a human reader's error recovery heuristics include checking previous words that caused processing difficulty and that this heuristic might utilize a memory trace of previous word-sense decisions, though this is not the only interpretation they offer. Thus, while ATLAST differs in many ways from the model of Carpenter and Daneman, especially in regard to the issue of reprocessing the input text, the latter model at least recognizes the plausibility of the principle of retention in explaining a reader's ability to recover from incorrect inferences made while reading misleading text.

V. Conclusion

The principle of retaining rejected inference paths within the larger framework of a relational network provides a simple but effective mechanism for recovering from erroneous inferences during text understanding, but only if there is a way to locate and re-evaluate the retained paths at the appropriate times.

From a practical perspective, the principle of inference retention could be incorporated into new or existing text understanding systems in order to enable them to correct erroneous decisions. From a cognitive modeling perspective, however, the jury is still out on the issue of inference retention. While a model like ATLAST demonstrates the plausibility of the theory, only psycholinguistic experiments designed specifically to test for retention will be able to confirm or deny the validity of the theory.

References

Carpenter, P.A., & Daneman, M. (1981). Lexical retrieval and error recovery in reading: A model based on eye fixations. *Journal of Verbal Learning and Verbal Behavior*, *20*, 137-160.

Charniak, E. (1983). Passing markers: A theory of contextual influence in language comprehension. *Cognitive Science*, *7*, 171-190.

Cottrell, G.W. (1984). A model of lexical access of ambiguous words. *Proceedings of the National Conference on Artificial Intelligence*, Austin, TX.

Granger, R.H. (1980). When expectation fails: Towards a self-correcting inference system. *Proceedings of the First Annual National Conference on Artificial Intelligence*, Stanford, CA.

Granger, R.H., & Holbrook, J.K. (1983). Perseverers, recencies, and deferrers: New experimental evidence for multiple inference strategies in understanding. *Proceedings of the Fifth Annual Conference of the Cognitive Science Society*, Rochester, NY.

Granger, R.H., Holbrook, J.K., & Eiselt, K.P. (1984). Interaction effects between word-level and text-level inferences: On-line processing of ambiguous words in context. *Proceedings of the Sixth Annual Conference of the Cognitive Science Society*, Boulder, CO.

Hirst, G. (1984). Jumping to conclusions: Psychological reality and unreality in a word disambiguation program. *Proceedings of the Sixth Annual Conference of the Cognitive Science Society*, Boulder, CO.

Holbrook, J.K., Eiselt, K.P., Granger, R.H., & Matthei, E.H. (1987). (Almost) never letting go: Inference retention during text understanding. In S.L. Small, G.W. Cottrell, & M.K. Tanenhaus (Eds.), *Lexical ambiguity resolution in the comprehension of human language*. Los Altos, CA: Morgan Kaufmann (to appear).

Hudson, S.B., & Tanenhaus, M.K. (1984). Ambiguity resolution in the absence of contextual bias. *Proceedings of the Sixth Annual Conference of the Cognitive Science Society*, Boulder, CO.

Norvig, P. (1983). Six problems for story understanders. *Proceedings of the National Conference on Artificial Intelligence*, Washington, DC.

O'Rorke, P. (1983). Reasons for beliefs in understanding: Applications of non-monotonic dependencies to story processing. *Proceedings of the National Conference on Artificial Intelligence*, Washington, DC.

Quillian, M.R. (1969). The teachable language comprehender: A simulation program and theory of language. *Communications of the ACM*, *12*(8), 459-476.

Riesbeck, C.K., & Martin, C.E. (1986). Direct memory access parsing. In J.L. Kolodner & C.K. Riesbeck (Eds.), *Experience, memory, and reasoning*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Seidenberg, M.S., Tanenhaus, M.K., Leiman, J.M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, *14*, 489-537.

Waltz, D.L., & Pollack, J.B. (1985). Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science*, *9*(1), 51-74.