# Jumping to Conclusions: Psychological Reality and Unreality in a Word Disambiguation Program

Graeme Hirst

Department of Computer Science
University of Toronto
Toronto, Canada M5S 1A4

## 1. Introduction

Human language understanding sometimes jumps to conclusions without having all the information it needs or even using all that it has. So, therefore, should any psychologically real language-understanding program. How this can be done in a discrete computational model is not obvious. In this paper, I look at three aspects of the problem:

- When is information ignored?
- When is a decision made out of impatience?
- When is no decision made at all?

I give illustrations of these problems in the domain of word disambiguation with the Polaroid Words system.

## 2. Polaroid Words

Polaroid[1] Words are a system for the disambiguation of words and case slots; they are a part of the Absity natural language understanding system (Hirst 1983a, 1983b). Their design is based in part on the results of recent psycholinguistic studies of word disambiguation that show that usually all meanings of an ambiguous word are activated and one is then chosen (Swinney 1979, Onifer and Swinney 1981, Seidenberg, Tanenhaus, Leiman and Bienkowski 1982). Thus in *The man walked on the deck*, both meanings of *deck*, 'pack of cards' and 'part of a boat', are activated below conscious awareness. This is in contrast to script-based models (Schank and Abelson 1977), in which the script acts as a context to predetermine a unique meaning for each ambiguous word, an approach clearly inadequate for polysemous words.

Each Polaroid Word (PW) is an independent procedure that is responsible for the disambiguation of one word or case slot in the input sentence. The PWs operate in parallel with one another and with other processes in the system.[2] There is one type of PW for nouns, another for prepositions, and so on. Each begins with a packet of knowledge that lists all possible meanings for its word or slot, and, as it obtains the knowledge to do so, eliminates all meanings that are inappropriate until just one is left. The possibilities in the PW's list are always well-formed semantic objects in the Frail frame system (Charniak, Gavin and Hendler 1983), and therefore may be used for retrieval and inference both by PWs and by other processes in the system, regardless of the extent to which disambiguation has or hasn't taken place. Polaroid Words and their many virtues are described more fully in Hirst (1983a) and Hirst and Charniak (1982); in the present paper we concentrate on their deficiencies.

## 3. Jumping to conclusions

### 3.1. Spreading activation and magic numbers

It has been shown that *semantic priming* by *spreading activation* (Collins and Loftus 1975) is important in human lexical disambiguation. Accessing a concept in memory temporarily

This work was carried out while I was at the Department of Computer Science, Brown University, Providence, Rhode Island. Financial support was provided in part by the U.S. Office of Naval Research under contract number N00014-79-C-0592. I am grateful to Eugene Charniak for discussions from which this work developed.

[1] *Polaroid* is a trademark of the Polaroid Corporation.

[2] In this respect, they bear a superficial similarity to Small's (1980) Word Experts.

activates both that concept and those closely connected to it, facilitating their subsequent retrieval. Seidenberg *et al* found that strong semantic priming of one sense of an ambiguous word was the only case where not all meanings were considered. For example, in ***The bridge player trumped the spade***, the word ***bridge*** primes the 'playing card suit' meaning of ***spade***, and 'digging instrument' is not considered.

To account for the effects of spreading activation in ambiguity resolution, Polaroid Words use *marker passing*, a discrete model of spreading activation in a network of frames and slots in the Frail representation. Marker passing can be thought of as spreading tokens along the arcs of the representation, marking each node reached, until all nodes within a certain distance of the origin have been marked. The trails of marks thus created are called *paths*. Markers may be passed along any connection in the network: from frame to slot, slot to filler, slot to constraint, class to sub-class, and so on. Markers are passed only to nodes within a few steps of the origin; otherwise, of course, the whole knowledge base would always get marked, a useless situation.

Before it does anything else, a PW checks whether one of its possibilities has, as result of previous activity, received a marker. If so, it decides immediately on this possibility without any further consideration. Otherwise, it asks Frail to start passing markers from each of its possibilities. If one of the paths so created intersects with a previously made path, this is taken as evidence that the origin is the appropriate sense of the ambiguous word. The closer the intersection is to the origin, the stronger the connection is considered to be; if the path is strong enough, the indicated sense is chosen. (If no such intersections are found, or only weak ones, the PW resorts to other methods, described in Hirst (1983a; Hirst and Charniak 1982).)[3] For example, in sentence (1):

(1) The plane taxied to the terminal.

the ambiguous words *plane* and *terminal* are resolved by finding the path between their aviation-related senses, but finding no path between any active concept and their other senses.

[3]Weak paths are not ignored entirely; rather, additional evidence is sought before a final decision is made.

The problem that immediately arises with this scheme is that of setting thresholds. How far from the origin should marker passing go? How strong does a path have to be before the PW can jump to a conclusion without considering other evidence? It is clear that there are psychologically real thresholds, for they sometimes result in people misinterpreting *negatively primed* ambiguities:[4]

(2) The astronomer married the star.

(3) The rabbi was hit on the temple.

(4) The sailor ate the submarine.

(5) The catcher filled the pitcher.

Although the selectional restrictions on *marry* in (2) are sufficient to uniquely determine the sense of *star* as 'celebrity', spreading activation from the meaning of *astronomer* causes most listeners to select the sense 'celestial object', despite the nonsensical result.[5] That is, the human disambiguation mechanism will sometimes wrongly jump to a conclusion — and PWs are likewise fooled by these sentences — even though information is present that would let it avoid the error. On the other hand, people generally have no trouble with the following sentences, which fall on the other side of the thresholds:[6]

(6) The lawyer bent the bar.
(*bar* ≠ 'legal profession')

(7) The dog chewed the bark.
(*bark* ≠ 'dog noise')

(8) The statistician sat on the table.
(*table* ≠ 'array of figures')

At present in Polaroid Words, markers are passed to nodes up to four steps away from the origin; but this threshold is just a magic number chosen arbitrarily, and is dependent upon the exact degree of coarseness of the Frail knowledge representation. What we need in order to determine a more realistic threshold is a large set of

[4]Sentences (3) and (5) are from Reder (1983).

[5]In the other sentences, *temple* = 'part of skull'; *submarine* = 'sandwich'; *pitcher* = 'jug'.

[6]While individuals vary as to exactly which sentences fall where, a disambiguation system with claims to psychological reality should be in accord with the general consensus.

data on the subjective semantic distance between many different concepts, and then see how this translates into "physical distance" in Frail (or other representation of choice). Word association norms (*e.g.*, Postman and Keppel 1970) may provide an initial base for such a set of data.

Getting the thresholds right in marker passing is important not only so that Polaroid Words can confidently use marker passing as a disambiguation cue, but also because properly deployed marker passing has many other uses in cognitive modeling; these include context determination and explanation finding (Charniak 1983).

### 3.2. Impatience

Recent psycholinguistic research (*e.g.*, Marslen-Wilson and Tyler 1980) has emphasized human language understanding's following the principle of "do it as early as possible" — that interpretation happens as soon as sufficient information is available, and the interpretation of earlier parts of a sentence is used to guide the interpretation of the later parts. This principle is followed by Absity (Hirst 1983a, 1983b), the system of which Polaroid Words form a part.

There are very few data, however, on how quickly lexical disambiguation takes place in humans. Almost all studies of disambiguation have only considered the special case in which sufficient disambiguation information is present when the ambiguous word occurs; often, the test word is the last of the sentence. Under these conditions, disambiguation is extremely rapid — between 100 and 200 msec (Lucas 1983).

But what of cases in which the necessary information is not initially present? How long will people wait for it before jumping to a conclusion with partial information? The following examples are both processed without error, although the final noun phrase has to be interpreted before *book* can be disambiguated as 'literary work' or 'printed volume':

(9) Nadia's favorite book is *The House at Pooh Corner*.

(10) Nadia's favorite book is her signed first edition of *The House at Pooh Corner*.

Thus, in at least some cases people will wait until the end of the clause. On the other hand, it is my intuition that *fans* is disambiguated as 'devotee' in (11) as soon as the verb *lined up* is processed:

(11) The fans were lined up for hours to buy the Stones tickets.

even though one can construct quite reasonable (albeit less probable) sentences that start the same way and in which *fan* means 'air-moving device':

(12) The fans were lined up awaiting their final factory inspection.

This suggests that PWs should use a *cumulating evidence* approach and jettison unlikely alternatives quickly if there is no positive evidence for them. That is, one does not make an immediate best guess, but one does make a reasonable guess as soon as there is enough information to do so, even if one cannot be definite. This has the advantage of helping to prevent combinatorial explosion.

However, I have been loath to consider using this approach in Polaroid Words, in view of the dearth of data on the corresponding human behavior and the fuzziness of the whole notion. Any interim solution would have to fall back on the magic numbers we have already bemoaned. Nevertheless, PWs do use the relative frequency of the various meanings of an ambiguous word in some of their decisions, but since we know little yet of how people use frequencies in disambiguation (see Hirst 1983a) we have limited their use in PWs to tidying up loose ends at the end of a sentence. Another possibility might be to add a mechanism that watches out for looming combinatorial explosion and forces PWs to make an early guess if it senses danger. (In Hirst 1983a, I discuss how the demands of structural disambiguation may force PWs to make an early decision, also in order to avoid combinatorial explosion.)

### 3.3. Cowardice

Despite everything we have said above, it is obvious that some sentences are genuinely ambiguous to people. It is therefore inappropriate for a disambiguation process to jump to a conclusion in such cases or to take extraordinary measures or go to heroic efforts to resolve residual problems. That is, PWs should be afraid to jump to a conclusion if the leap seems too great. If reasonable efforts fail, they can always ask the user what he or she really meant:

(13) USER: I need some information on getting rid of moles.
SYSTEM: Are you troubled by unsightly blemishes, by those lovable but destructive insectivorous garden pests, by uterine debris, or by enemy secret agents that have penetrated deep into your organization?

(PWs do not actually have such a natural language response component.)

## 4. Conclusion

The notion of jumping to a conclusion when there is "enough" evidence is an inherently fuzzy one, but one that is clearly involved in word disambiguation, as well as other cognitive processes. The "easy" solution, using magic numbers in a delicately balanced knowledge base, is obviously inadequate. A better understanding of the time course of human word disambiguation is needed before the psychological reality of Polaroid Words can be improved.

## References

CHARNIAK, Eugene (1983). "Passing markers: A theory of contextual influence in language comprehension." *Cognitive science*, 7(3), July-September 1983, 171-190.

CHARNIAK, Eugene, GAVIN, Michael and HENDLER, James (1983). "The Frail/NASL reference manual." Technical report CS-83-06, Department of Computer Science, Brown University, February 1983.

COLLINS, Allan M and LOFTUS, Elizabeth F (1975). "A spreading-activation theory of semantic processing." *Psychological review*, 82(6), November 1975, 407-428.

HIRST, Graeme (1983a). *Semantic interpretation against ambiguity*. Doctoral dissertation (technical report CS-83-25), Department of Computer Science, Brown University, December 1983.

HIRST, Graeme (1983b). "A foundation for semantic interpretation." *Proceedings, 21st Annual Meeting of the Association for Computational Linguistics*, Cambridge, Massachusetts, June 1983. 64-73. (Also available as technical report CS-83-03, Department of Computer Science, Brown University, January 1983.)

HIRST, Graeme and CHARNIAK, Eugene (1982). "Word sense and case slot disambiguation." *Proceedings, National Conference on Artificial Intelligence (AAAI-82)*, Pittsburgh, August 1982. 95-98.

LUCAS, Margery (1983). "Lexical access during sentence comprehension: Frequency and context effects." *Proceedings, Fifth Annual Conference of the Cognitive Science Society*, Rochester, New York, May 1983.

MARSLEN-WILSON, William D and TYLER, Lorraine Komisarjevsky (1980). "The temporal structure of spoken language understanding." *Cognition*, 8(1), March 1980, 1-71.

ONIFER, William and SWINNEY, David (1981). "Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias." *Memory and cognition*, 9(3), May 1981, 225-236.

POSTMAN, Leo and KEPPEL, Geoffrey (editors) (1970). *Norms of word association*. New York: Academic Press, 1970.

REDER, Lynne M (1983). "What kind of pitcher can a catcher fill? Effects of priming in sentence comprehension." *Journal of verbal learning and verbal behavior*, 22(3), April 1983, 189-202.

SCHANK, Roger and ABELSON, Robert (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1977.

SEIDENBERG, Mark S, TANENHAUS, Michael K, LEIMAN James M and BIENKOWSKI, Marie A (1982). "Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing." *Cognitive psychology*, 14(4), October 1982, 489-537.

SMALL, Steven L (1980). *Word expert parsing: A theory of distributed word-based natural language understanding*. Doctoral dissertation (technical report 954), Department of Computer Science, University of Maryland, September 1980.

SWINNEY, David (1979). "Lexical access during sentence comprehension: (Re)Consideration of context effects." *Journal of verbal learning and verbal behavior*, 18(6), December 1979, 645-659.