A SEMANTIC PROCESS FOR SYNTACTIC DISAMBIGUATION

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

ABSTRACT

Structural ambiguity in a sentence cannot be resolved without semantic help. We present a process for structural disambiguation that uses verb expectations, presupposition satisfaction, and plausibility, and an algorithm for making the final choice when these cues give conflicting information. The process, called the Semantic Enquiry Desk, is part of a semantic interpreter that makes sure all its partial results are well-formed semantic objects; it is from this that it gains much of its power.

1. INTRODUCTION

It is universally accepted that syntactic analysis of natural language requires much semantic knowledge, and it is generally accepted that semantic analysis requires much syntactic knowledge. (Convincing arguments for the latter are presented by Marcus 1984.) The goal of the present research is a system in which syntax and semantics relate well to one another, and are both properly deployed to find the semantic interpretation of the input, dealing with ambiguities of word sense, case slot filling, and syntactic structure.

We are assuming a frame-like representation of knowledge with a suitable retrieval and inference engine — in particular, we are using the FRAIL frame system (Charniak, Gavin and Hendler 1983). In Hirst 1983a, 1983b, we showed how such a representation can provide an adequate notion of "semantic object", in the Montague (1973) sense, and developed a system named Absity in which semantic rules operated in tandem with corresponding syntax rules upon corresponding objects. The system has some of the flavor of Montague's, but replaces possible worlds with A.I.-style representations and the categorial grammar with an A.I.-style parser with wider syntactic coverage.

A mechanism for word sense and case slot disambiguation that worked in conjunction with Absity was presented by Hirst and Charniak (1982; Hirst 1983a). This mechanism, called Polaroid Words, drew much of its power from the design of Absity, which ensured that all semantic entities in the system were always well-formed semantic objects in the FRAIL representation and inference system. It remained, however, to deal with ambiguities of syntactic structure. We now present a mechanism for this, the Semantic Enquiry Desk (*SED*). There are many types of structural ambiguity (see Hirst 1983a for a long list); the SED handles two important kinds — prepositional phrase attachment and problems of gap-finding in relative clauses — and provides a foundation for the development of methods for dealing with other kinds. In this paper, we will look at prepositional phrase (*PP*) attachment, in which a PP may be attached to either the verb phrase (*VP*) of the clause as a case slot filler, or to a noun phrase (*NP*) as a modifier.

We are using a parser similar to Marcus's (1980) limited-lookahead deterministic parser, Parsifal. Our approach could, however, be adapted to other types of parser, provided only that they are able to give the SED sufficient information.

2. TWO THEORIES OF STRUCTURAL DISAMBIGUATION

The SED synthesizes two rather different theories of structural disambiguation: The lexical preference theory of Ford, Bresnan and Kaplan (1982) and the presupposition minimization theory of Crain and Steedman (1984). We explain each briefly.

Ford, Bresnan and Kaplan (*FBK*) show that disambiguation strategies such as Minimal Attachment (Frazier and Fodor 1978) that are based solely on syntactic preferences are inadequate to account for the resolution preferences that people exhibit in experiments. Rather, the preferred structure can change with the verb:

(1) The women discussed the dogs on the beach.
(i.e., NP attachment: The dogs on the beach were discussed by the women.)

(2) The women kept the dogs on the beach.
(i.e., VP attachment: On the beach was where the women kept the dogs.)

FBK propose a theory of *lexical preferences*, in which each verb is marked with the cases that are generally used with it. Each PP is assumed to be one of these *expected cases*, to be attached to the VP, and is interpreted as such if at all possible, until the last expected

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. Preparation of this paper was supported by grants from the Connaught Fund, University of Toronto, and the Natural Sciences and Engineering Research Council of Canada.

case is filled; subsequent PPs are assumed to be NP modifiers of the final expected case. These assumptions are dropped if an anomalous interpretation would result, or if pragmatics overrule them. FBK show that this principle accounts for some other kinds of structural ambiguity as well as PP attachment.

A very different theory of structural disambiguation has been proposed by Crain and Steedman (1984), who claim that discourse context and, in particular, *presupposition* and *plausibility*, are paramount in structural disambiguation. The presuppositions of a sentence are the facts that a sentence assumes to be true and the entities that it assumes to exist. If a sentence presupposes information that the reader does not have, she has to detect and invoke these *unsatisfied* presuppositions. People have no trouble doing this, though there is evidence that it increases comprehension time (Haviland and Clark 1974); Weischedel (1979) has shown how presuppositions may be determined as the sentence is parsed.

Crain and Steedman hypothesize *The Principle of Parsimony:* the reading that leaves the fewest presuppositions unsatisfied is the one to be favored, other things being equal. This is a particular case of the *Principle of A Priori Plausibility:* prefer the reading that is more plausible with regard to either general knowledge about the world or specific knowledge about the universe of discourse, other things being equal. These principles can explain well-known garden-path sentences such as (3):

(3) The horse raced past the barn fell.

The correct parse presupposes both the existence of a particular horse and that this horse is known to have raced past a barn, presuppositions unsatisfied in the null context. The incorrect parse, the one that garden-paths, only presupposes the first of these; the other is taken as new information that the sentence is conveying. The Principle of Parsimony claims that the garden-path parse is chosen just because it makes fewer unsatisfied presuppositions. Experiments by Crain and Steedman support this analysis, and suggest that Ford, Bresnan and Kaplan's results are just artifacts of their use of the null context, not controlling for unsatisfied presuppositions. Nevertheless, FBK's experiments found ambiguities whose preferred resolutions do seem to require an explanation in terms of lexical preference rather than presupposition or plausibility (Hirst 1983a). A more detailed discussion of the two approaches may be found in Hirst 1983a.

3. PREPOSITIONAL PHRASE ATTACHMENT

Many easy cases of prepositional phrase attachment can be handled by simple and absolute lexical and syntactic knowledge about allowed attachment. For example, few verbs will admit the attachment of a PP whose preposition is *of*, and such knowledge may be included in the lexical entry for each verb.

For those cases where deeper consideration is necessary, the SED's approach to PP attachment is to synthesize the two approaches of the previous section. There are four things needed for this:

- An annotation on each verb sense as to which of its cases are "expected".
- A method for determining the presuppositions that would be engendered by a particular PP attachment, and for testing whether they are satisfied or not.
- A method for deciding on the relative plausibility of a PP attachment.
- A method for resolving the matter when the preceding strategies give contradictory recommendations.

3.1. Verb annotations

The first requirement, annotating verbs for what they expect, is straightforward once we have data on verb preferences. These data should come from formal experiments on people's preferences, such as the one Ford, Bresnan and Kaplan (1982) ran, or from textual analysis; however, for a small, experimental system such as ours, the intuitions of the author and his friendly informants will suffice. We classify cases as either **compulsory**, **preferred**, or **unlikely**.

3.2. Testing for presupposition satisfaction

The next requirement is a method for deciding whether a particular PP attachment would result in an unsatisfied presupposition. Now, there is a simple trick, first used by Winograd (1972), for determining many PP attachments: try each possibility and see if it describes something that is known to exist. For example, sentence (4):

(4) Put the block in the box on the table.

could be asking that **the block** be placed in **the box on the table**, or that **the block in the box** be placed on **the table**. The first reading can be rejected if *the block* does not in context uniquely identify a particular block, or if there is no box on the table, or if *the box on the table* does not uniquely identify a particular box. Similar considerations may be applied to the second reading. (If neither reading is rejected, or if both are, the sentence is ambiguous, and Winograd's program would seek clarification from the user.) Crain and Steedman have called this technique *The Principle of Referential Success*: a reading that succeeds in referring to an entity already established in the hearer's mental model of the domain of the discourse is favored over one that does not.

We will show that the Principle of Referential Success suffices in checking for unsatisfied presuppositions. We observe the following.[1] First, a definite non-generic NP presupposes that the thing it describes exists and is available in the focus or knowledge base for felicitous (unique) reference; an indefinite NP presupposes only the plausibility of what it describes. Thus, *a blue chip-*

munk presupposes only that the concept of a **blue chipmunk** is plausible; *the blue chipmunk* further presupposes that there is exactly one blue chipmunk available for ready reference. Second, the attachment of a PP to an NP results in new presuppositions for the new NP thus created, but cancels any uniqueness aspect of the referential presuppositions of both its constituent NPs. Thus, *the red tree with the blue chipmunk* presupposes that there is just one such tree available for reference (and that such a thing is plausible); the plausibility and existence of a **red tree** and a **blue chipmunk** continue to be presupposed, but their uniqueness is no longer required. Third, the attachment of a PP to a VP creates no new presuppositions, but rather always indicates new (unpresupposed) information.[2]

These observations allow us to "factor out" most of the presupposition testing: the candidate attachments will always score equally for unsatisfied presuppositions, except that VP attachment wins if the NP candidate is definite but NP attachment would result in reference to an unknown entity. On the other hand, if NP attachment would result in a felicitous definite reference, the number of unsatisfied presuppositions will remain the same for both attachments, but by the Principle of Referential Success we will prefer the NP attachment.[3]

Testing for this is easy for the SED because of the property of Absity that the semantic objects associated with the syntactic constituents are all well-formed FRAIL objects. The SED puts them into a call to FRAIL to see whether the mooted NP-attachment entity exists in the knowledge base or not. (The entity may be there explicitly, or its existence may be inferred; that is up to FRAIL.)

If the entity is found, the presupposition is satisfied, and the PP should be attached to the NP; otherwise, if the presupposition is unsatisfied, or if no presupposition was made, the VP is favored for the PP.

As an example, let's suppose the SED needs to decide on the attachment of the PP in (5):

(5) Ross saw the man with the telescope.

It will have the semantic objects for *see*, *the man*, and *with the telescope*, the last having two possibilities, one for each attachment mooted. It constructs the FRAIL statement (6) for the NP attachment:

(6) **(the ?x (man ?x (attr=(the ?y (telescope ?y)))))**

[1]The proof of the generality of these observations is by absence of counterexample. If the reader has a counterexample, she should notify me promptly.

[2]This is not quite true; sentences asserting a change of state presuppose that the new state did not previously hold.

[3]A corollary of this is that a PP is never attached to an indefinite NP if VP attachment is at all possible, except if the NP is the final expected argument. This seems too strong, and our rule will probably need toning down. This corollary is not completely out of line, however, as definiteness does influence attachment; see Hirst 1983a.

If this returns an instance, **man349** say, then the SED knows that presupposition considerations favor NP attachment; if it returns **nil**, then it knows they favor VP attachment.

3.3. Plausibility

Now let's consider the use of plausibility to evaluate the possible PP attachments. In the most general case, deciding whether something is plausible is extremely difficult, and we make no claims to having solved the problem. In the best of all possible worlds, FRAIL would be able to answer most questions on plausibility, and the slot restriction predicates on frames would be *defined* to guarantee plausibility; but, of course, we don't know how to do that.

However, there are two easy methods of testing plausibility that we can use that, though non-definitive, will suffice in many cases. The first of these, used in many previous systems, is selectional restrictions. In the present system, these are applied as *slot restriction predicates* by the case slot disambiguation part of Polaroid Words even before the SED becomes involved, and are often adequate by themselves. While satisfying the predicates does not guarantee plausibility, failing the predicates indicates almost certain implausibility.

The second method is what we shall call the *Exemplar Principle* (a weak form of the Principle of Referential Success): an object or action should be considered plausible if the knowledge base contains an instance of such an object or action, or an instance of something similar. Again, the SED can easily construct from the semantic objects supplied to it the FRAIL call to determine this. For example, if the SED wants to test the plausibility of **a cake with candles** or **operate with a slug**, it looks in the knowledge base to see if it has run across such a thing before:

(7) **(a ?x (cake ?x (attr=(some ?y (candle ?y)))))**

(8) **(a ?x (operate ?x (instrument=(a ?y (slug ?y)))))**

If it finds an instance, it takes the attachment to be plausible. If no such item is found, the matter is unresolved.[4] Thus the results of plausibility testing by the SED will be either **exemplar exists** or **can't tell**.[5]

3.4. Making the attachment decision

The SED's last requirement is a method for deciding on the PP attachment, given the results of verb expectation and presupposition and plausibility testing. If all

[4]Various recovery strategies suggest themselves; see Hirst 1983a.

[5]With a large knowledge base it may be possible to assign ratings based on the number of exemplars found; an item that has a hundred exemplars would be considered more plausible than one with only one exemplar, other things being equal. See Hirst 1983a for discussion.

TABLE 1.

DECISION ALGORITHM FOR RESTRICTIVE PP ATTACHMENT (ONE VP AND ONE NP)

[Referential success]
if NP attachment gives referential success
then attach to NP

[Plausibility]
else if an exemplar is found for exactly one attachment
then make that attachment

[Verb expectations]
else if verb expects a case that the preposition could be flagging
then attach to VP
else if the last expected case is open
then attach to NP

[Avoid failure of reference]
else if NP attachment makes unsuccessful reference
then attach to VP
else sentence is ambiguous, but prefer VP attachment anyway.

agree on how the attachment should be made, then everything is fine. However, as Ford, Bresnan and Kaplan (1982) make clear, verb expectations are only biases, not absolutes, and can be overridden by conflicting context and pragmatic considerations. Therefore, the SED needs to know when overriding should occur. Table 1 shows a decision algorithm for this that assumes that one VP and one NP are available for attaching the PP to. (An algorithm for the case of several available NPs is presented in Hirst 1983a.) The algorithm gives priority to ruling out implausible readings, and favors NP attachments that give referential success (referential success is tried first, since it is a stronger condition); if these tests don't resolve matters, it tries to use verb expectations.[6] If these don't help either, it goes for VP attachment (*i.e.*, Minimal Attachment), since that is where structural biases seem to lie, but it is more confident in its result if an unsatisfied presupposition contraindicates NP attachment.

Some sentences for which the algorithm gives the correct answer are shown in Table 2. We also show a couple of sentences on which the algorithm fails. The fault in these cases seems to be not in the algorithm but rather in the system's inability to use world knowledge as well as people do. I can't believe that people have some sophisticated mental algorithm that tells them how to attach PPs in those awkward cases where several different possibilities all rate approximately the same; rather, they use a simple algorithm and lots of knowledge, and in the rare awkward (and, probably, artificial) case, either ask for clarification,

[6]There are sentences in which verb expectations prevail over plausibility; see Hirst 1983a. Ideally, the SED would react to these sentences the way people do; however, the procedure we present errs on the side of common sense.

TABLE 2.

PPs THAT ARE AND AREN'T CORRECTLY ATTACHED

PPs THAT ARE CORRECTLY ATTACHED

The women discussed the dogs on the beach.
NP-attached.

The women discussed the tigers on the beach.
NP-attached if there are tigers on the beach, but VP-attached if no examples of tigers on the beach *are found.*

Ross bought the book for Nadia.
VP-attached unless there is a book for Nadia available for reference.

Ross included the book for Nadia.
NP-attached, as per FBK's preference data.

PPs THAT ARE NOT CORRECTLY ATTACHED

The women discussed dogs on the beach.
NP-attached because dogs on the beach *is plausible and doesn't fail referentially, though VP attachment seems to be preferred by informants.*

The women discussed the dogs at breakfast.
NP-attached like the dogs on the beach, *because the subtle unusualness of* the dogs at breakfast *is not detected.*

choose an attachment almost at random, or use conscious higher-level inference (perhaps the kind used when trying to figure out garden paths) to work out what is meant.

4. MUFFLING COMBINATORIAL EXPLOSIONS

The preceding discussion assumed that while the meaning of the preposition of the PP may be unresolved, the potential attachment heads (*i.e.*, the noun of the NP and the verb of the VP) and the remainder of the PP were all either lexically unambiguous or already disambiguated. Now let's consider what happens if they are not, that is, if the words that must be used by the SED to decide on an attachment are ambiguous. We will see that the SED's decision will often as a side effect allow the words to be disambiguated as well.

In principle, the number of combinations of meanings of the words that are not yet disambiguated could be large. For example, if the two potential attachment heads, the preposition, and the prepositional complement all have three uneliminated senses, then 81 (*i.e.*, 3^4) combinations of meanings could be constructed. In practice, however, many combinations will not be semantically possible, as one choice will constrain another — the choice for the verb will restrict the choices for the nouns, for example. Moreover, such multiple ambiguities are probably extremely rare. (I was unable to construct an example that didn't sound artificial.) It is my intuition that verbs are almost always disambiguated by the NP or PP that immediately follows them, before any PP attachment questions can

arise. Moreover, the SED could use the strategy that if the verb remains ambiguous when PP attachment is being considered and combinatorial explosion seems imminent, the verb is *required* by the SED to disambiguate itself forthwith, even if it has to guess.[7] (This is in accord with Just and Carpenter's (1980) model of reading, in which combinatorial explosion is avoided by judiciously early choice of word senses.)

Given, then, a manageably small number of lexical ambiguity combinations, structural disambiguation by the SED may proceed as before. Now, however, each attachment must be tried for each combination. The type of attachment that scores best for some combination is then chosen, thereby also choosing that combination as the resolution of the lexical ambiguities. For example, if combination *A* suggests NP attachment on the basis of referential success, thus beating combination *B*'s suggestion of VP attachment on the basis of plausibility, then both NP attachment and the word senses in combination *A* are declared winners. Ties are, of course, possible, and may indicate genuine ambiguity; see Hirst 1983a for discussion.

5. OTHER STRUCTURAL AMBIGUITIES

In Hirst 1983a, I show how similar techniques may be used for gap-finding in relative clauses, and give some preliminary suggestions on how the SED may also handle particle detection, relative clause attachment, and adverb attachment.

6. CONCLUSION

Like Polaroid Words, the Semantic Enquiry Desk gains much of its power from the property of Absity that its partial results, the constituents with which the SED works, are always well-formed FRAIL objects, enabling it to use the full power of a frame and inference system. Even if the correct choice of object for an ambiguous word is not yet known, the alternatives will be well-formed and easily accessible.

ACKNOWLEDGEMENTS

I am grateful to Eugene Charniak for discussions from which this work developed, to Stephen Crain for information about his work, and to Nadia Talent and Yawar Ali for their comments upon an earlier draft of the paper.

REFERENCES

CHARNIAK, Eugene; GAVIN, Michael Kevin and HENDLER, James Alexander (1983). "The FRAIL/NASL reference manual." Technical report CS-83-06, Department of Computer Science, Brown University, Providence, RI 02912. February 1983.

CRAIN, Stephen and STEEDMAN, Mark (1984). "On not being led up the garden path: The use of context by the psychological parser." in DOWTY, David R; KARTTUNEN, Lauri Juhani and ZWICKY, Arnold M (editors). *Syntactic theory and how people parse sentences* (= Studies in natural language processing 1). Cambridge University Press, 1984.

FORD, Marilyn; BRESNAN, Joan Wanda and KAPLAN, Ronald M (1982). "A competence-based theory of syntactic closure." in BRESNAN, Joan Wanda (editor). *The mental representation of grammatical relations.* Cambridge, MA: The MIT Press, 1982. 727-796.

FRAZIER, Lyn and FODOR, Janet Dean (1978). "The sausage machine: A new two-stage parsing model." *Cognition*, **6**(4), December 1978, 291-325.

HAVILAND, Susan E and CLARK, Herbert H (1974). "What's new? Acquiring new information as a process in comprehension." *Journal of verbal learning and verbal behavior*, **13**(5), October 1974, 512-521.

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

HIRST, Graeme (1983b). "A foundation for semantic interpretation." [1] *Proceedings, 21st Annual Meeting of the Association for Computational Linguistics*, Cambridge, Massachusetts, June 1983. 64-73. [2] Technical report CS-83-03, Department of Computer Science, Brown University, Providence, RI 02912. 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.

JUST, Marcel Adam and CARPENTER, Patricia A (1980). "A theory of reading: From eye fixations to comprehension." *Psychological review*, **87**(4), July 1980, 329-354.

MARCUS, Mitchell P (1980). *A theory of syntactic recognition for natural language.* Cambridge, MA: The MIT Press, 1980.

MARCUS, Mitchell P (1984). "On some inadequate theories of human language processing." in BEVER, Thomas G; CARROLL, John M and MILLER, Lance A (editors). *Six distinguished lectures on language.* Cambridge, MA: The MIT Press [to appear].

MONTAGUE, Richard (1973). "The proper treatment of quantification in ordinary English." [1] in HINTIKKA, Kaarlo Jaakko Juhani; MORAVCSIK, Julius Matthew Emil and SUPPES, Patrick Colonel (editors). *Approaches to natural language: Proceedings of the 1970 Stanford workshop on grammar and semantics.* Dordrecht: D. Reidel, 1973. 221-242. [2] in: THOMASON, Richmond Hunt (editor). *Formal philosophy: Selected papers of Richard Montague.* New Haven: Yale University Press, 1974. 247-270.

WEISCHEDEL, Ralph M (1979). "A new semantic computation while parsing: Presupposition and entailment." in OH, Choon-Kyu and DINNEEN, David A (editors). *Syntax and semantics 11: Presupposition.* New York: Academic Press, 1979. 155-182.

WINOGRAD, Terry (1972). [1] "Understanding natural language." *Cognitive Psychology*, **3**(1), 1972, 1-191. [2] *Understanding natural language.* New York: Academic Press, 1972.

[7] This strategy is not yet implemented in the SED.