Chapter 3

TRADITIONAL APPROACHES TO ANAPHORA

They went about and sang of Rāma's deeds; and Rāma heard of it, and he called an assembly of the Brāhmans and all kinds of grammarians . . . and the hermit children sang before them all.

— The Rāmāyana [1]

In this chapter and Chapter 5 I describe and evaluate some of the approaches that have been taken to anaphora, with respect to NLU systems, over the past years. I have divided them very roughly into two classes: traditional and modern. The traditional systems tend not to recognize as a separate problem the question of what is or isn't in consciousness. Rather, they assume that, other things being equal, the set of possible referents is exactly the set of NPs (or whatever), from the whole of the preceding text, in strict order of recency. Their resolution methods tend to work at the sentence level, and may bring to bear world knowledge and low-level linguistic knowledge. Antecedents not explicit in the text are not handled. This characterization is of course a generalization; not all approaches classified as traditional fit this description in every detail. On the other hand, modern methods recognize the importance of focus and discourse-level knowledge for resolution. Implicit antecedents may also be handled.

In this chapter, I review the traditional methods; in Chapter 5, the modern methods are presented.

3.1. Some traditional systems

First we will look at some of the systems that employed traditional anaphor resolution methods.

---

[1]From the translation in: Coomaraswamy, Ananda K and The Sister Nivedita of Rāmakrishna–Vivekānanda (Margaret E Noble). *Myths of the Hindus & Buddhists*. [1] Harrap, 1913. [2] New York: Dover, 1967. page 110.

> *I lisp'd in Numbers.*
> — *Alexander Pope*[2]

3.1.1. STUDENT

The high-school algebra problem answering system STUDENT (Bobrow 1964), an early system with natural language input, has only a few limited heuristics for resolving anaphors and, more particularly, anaphor-like paraphrases and incomplete repetitions. For example, in a question such as (3-1):

> (3-1) The number of soldiers the Russians have is half the number of guns they have. The number of guns is 7000. What is the number of soldiers they have?

the system will first try to solve the problem treating *the number of soldiers the Russians have* and *the number of soldiers they have* as two separate and distinct variables. Upon failure, it will eventually identify the two phrases by noting that they are identical up to the pronoun in the second. Similarly, it will identify *the number of guns* with *the number of guns they have* by the fact that the former is contained in and occurs after the latter. STUDENT does not actually resolve the pronouns at all. Phrases containing *this* are usually taken to refer to the consequence of the immediately preceding item without looking at the rest of the phrase. Thus, in (3-2):

> (3-2) A number is multiplied by 6. This product is increased by 44.

the word *product* could be changed to *result* or *sasquatch* without changing the assumed referent of *this*. Cases like (3-3):

> (3-3) The price of a radio is 69.70 dollars. <u>This price</u> is 15% less than the marked price.

are apparently resolved through the two occurrences of the word *price*.

Clearly, these simple heuristics are easily fooled since the sentence is not even parsed in any real sense. For example, in (3-4) the two references to sailors would not be matched up, although modifications to the heuristics may change this:

> (3-4) The number of soldiers the Russians have is twice the number of sailors they have. The number of soldiers is 7000. How many sailors do the Russians have?

---

[2]From: An epistle to Dr Arbuthnot. 2 January 1735, line 128. in, inter alia: Pope, Alexander. *Imitations of Horace with an epistle to Dr Arbuthnot and the Epilogue to the Satires*. (= The Twickenham edition of the poems of Alexander Pope 4). London: Methuen, 1939.

However a sophisticated paraphrase of (3-4) would stand no chance:

(3-5) If the Russians have twice as many soldiers as sailors, and they have 7000 soldiers, how many sailors are there?

"No, no", said Anne. "That won't do. You must do something more than that."
"But what? All the good jobs are taken, and all I can do is lisp in numbers."
"Well, then, you must lisp", concluded Anne.
    — Aldous Leonard Huxley[3]

3.1.2. SHRDLU

Winograd's (1971, 1972) celebrated SHRDLU system employs heuristics much more complex than those of STUDENT, providing impressive and, for the most part, sophisticated handling of anaphors, including references to earlier parts of the conversation between the program and its user. The most important aspect of SHRDLU's handling of anaphors is that in checking previous noun groups as possible referents, it does not seize the first likely candidate for use, but rather checks all possibilities in the preceding text and assigns each a rating whereby the most plausible answer is selected. If none clearly stands out as a winner, the user is asked for help in choosing between the serious contenders.

Gross heuristics cover some simpler cases. If *it* or *they* occurs twice in the same sentence, or in two adjacent sentences, the occurrences are assumed to be coreferential. This usually works, but there are, as always, easy counterexamples, such as (3-6) (from Minsky 1968):

(3-6) He put the box on the table. Because it wasn't level, it slid off.

An anaphor which is part of its own referent, as (3-7):

(3-7) a block which is bigger than anything which supports it

can be detected and interpreted correctly by SHRDLU without infinite regression. Reference to events, as in (3-8):

(3-8) Why did you do it?

is resolved through always remembering the last event referred to.

---

[3]From: *Crome yellow.* New York: Harper, 1922.

Some contrastive uses of *one* can be handled, as in (3-9):

(3-9) a big green pyramid and a little <u>one</u>

A list of pairs of words like *big* and *little* that are often used contrastively is employed to work out that *little one* here means *little green pyramid* and not *little pyramid* or *little big green pyramid*. This method assumes no redundant information is given. Suppose your universe had three pyramids: a big blue one, a big green one and a little blue one. Then the above interpretation of (3-9) would have you looking for a little green pyramid which you don't have, when the speaker obviously meant the little blue one. Although the *big* in (3-9) is redundant and has resulted in an erroneous interpretation, it is a perfectly acceptable phrase which reflects the way people often talk.

The methods used for *one* are also used for incompletes that are cardinal numbers, such as in (3-10):

(3-10) Find the red blocks and stack up <u>three</u>.

3.1.3. LSNLIS

The Lunar Sciences Natural Language Information System (LSNLIS — also known as LUNAR) (Woods, Kaplan and Nash-Webber 1972; Woods 1977) uses an ATN parser (Woods 1970) and a semantic interpreter based on the principles of procedural semantics (Woods 1968).[4] It is in this latter component that the system resolves anaphoric references, giving full meaning to pronouns found in the parse tree.

The system distinguishes two classes of anaphors: PARTIAL and COMPLETE. A complete anaphor (of which there are three types) is a pronoun which refers to a complete antecedent noun phrase, while a partial one refers to only part of a preceding NP; that is, the first is an IRA and the second an ISA. (3-11) shows a complete anaphoric reference and (3-12) a partial one:

(3-11) Which coarse-grained rocks have been analyzed for cobalt? Which <u>ones</u> have been analyzed for strontium?

(3-12) Give me all analyses of sample 10046 for hydrogen. Give me <u>them</u> for oxygen.

Note that in (3-12), *them* refers to *all analyses of sample 10046*, whereas the NP in the antecedent sentence was *all analyses of sample 10046 for hydrogen*. Such partial anaphors are signalled by the presence of a relative clause or

---

[4]A useful overview of the whole LSNLIS system, together with a detailed critique of its anaphor handling capabilities, may be found in Nash-Webber (1976).

prepositional phrase modifying the pronoun; here it is *for oxygen*.

Partial anaphors are resolved by searching through antecedent noun phrases for one with a parallel syntactic and semantic structure. In (3-12), for example, the antecedent NP is found, and *for oxygen* substituted for *for hydrogen*. This method is not unlike Bobrow's in STUDENT (see section 3.1.1), but it works on the syntactic and semantic level rather than at the more superficial level of lexical matching with a little added syntax. It suffers however from the same basic limitation, namely that it can only resolve anaphors where the antecedent is of a similar structure. Neither (3-13) nor (3-14), for example, could have been used as the second sentence of (3-12):

(3-13) Give me the oxygen <u>ones</u>.

(3-14) Give me <u>those</u> that have been done for oxygen.

Three different methods are used for complete anaphoric references, the one chosen depending on the exact form of the anaphor. The first form includes a noun and uses the anaphor as a determiner:

(3-15) Do any breccias contain aluminium? What are <u>those</u> breccias?

The strategy used here is to search for a noun phrase whose head noun is *breccias*. Note that if the second sentence contained instead a paraphrase, such as *those samples*, this method would either find the wrong antecedent, or none at all, as there is no mechanism for recognizing the paraphrase.

The second form is a single pronoun:

(3-16) How much titanium is in type B rocks? How much silicon is in <u>them</u>?

In this case, more semantic information needs to be used. The semantic template which matches "ELEMENT BE IN" requires that the object of the verb be a SAMPLE, and this fact is used in searching for a suitable antecedent in this example. This is isomorphic to a weak use of a case-based approach (see sections 3.1.5 and 3.2.4).

The third type of complete anaphor is *one* and *ones*, as in (3-11). These are resolved either with or without modifiers like *too* and *also*. (Notice that if either of these modifiers were appended to (3-11), the meaning would be completely changed, the anaphor referring not to the first question but rather to its answer.) Resolution is by a method similar to that used for single pronouns.

The primary limitation of LSNLIS is that intrasentential anaphors cannot be resolved, because a noun phrase is not available as an antecedent until processing of the sentence containing it is complete.

3.1.3 LSNLIS

3.1.4. MARGIE and SAM

So far, the natural language systems based on conceptual dependency theory (Schank 1973), MARGIE (Schank, Goldman, Rieger and Riesbeck 1975; Schank 1975) and SAM (Schank and the Yale AI Project 1975; Schank and Abelson 1977; Nelson 1978), have apparently not been able to handle any form of anaphor much beyond knowing that *he* always refers to John (a pathetic victim of social brutalization) and *she* to Mary (a pathetic victim of John, who frequently beats and murders her).

However, the Conceptual Memory section of MARGIE (Rieger 1975) is able to resolve some limited forms of definite reference by inference. Conceptual Memory operates upon nonlinguistic representations of concepts based on Schank's conceptual dependency theory, and can perform sixteen types of inference, including motivational, normative, causative and resultative. For example, if the system knows of two people named Andy, one an adult and one an infant, it can work out which is the subject of (3-17):

(3-17) Andy's diaper is wet.

That conceptual dependency−based systems should be so limited with respect to reference is disappointing, as conceptual dependency may prove to be an excellent framework for inference on anaphors (see section 3.2.6).

3.1.5. A case-driven parser

In his case-driven parser, Taylor (1975; Taylor and Rosenberg 1975) uses case analysis (Fillmore 1968, 1977) to resolve anaphors.

Pronouns are only encountered by the parser when a particular verb case is being sought, thereby giving much information about its referent. Previous sentences and nonsubordinate clauses[5] are searched for a referent that fits the case and which passes other tests, usually SHOULD-BE and MUST-BE predicates, to ensure that it fits semantically. As the search becomes more desperate, the SHOULD-BE tests are relaxed. Locative and dummy-subject anaphors can also be resolved.

The parser will always take the first candidate that passes all the tests as the referent. This occasionally leads to problems, where there are two or more acceptable candidates, but the first one found is not the correct one.

---

[5]Subordinate clauses in English can contain anaphors, but Taylor's system will not find them.

How is this done? By fucking around with syntax.
— Tom Robbins[6]

3.1.6. Parse tree searching

An algorithm for searching a parse tree of a sentence to find the referent for a pronoun has been given by Jerry Hobbs (1976, 1977). The algorithm takes into consideration various syntactic constraints on pronominalization (see section 3.2.2) to search the tree in an optimal order such that the NP upon which it terminates is probably the antecedent of the pronoun at which the algorithm started. (For details of the algorithm, which is too long to give here, and an example of its use, see Hobbs (1976:8-13) or Hobbs (1977:2-7).)

Because the algorithm operates purely on the parse, it does not take into account the meaning of the text, nor can it find non-explicit antecedents. Nonetheless, Hobbs found that it gives the right answer a large proportion of the time.

To test the algorithm, Hobbs took text from an archaeology book, an Arthur Hailey novel and a copy of *Newsweek*. From each of these as much contiguous text as was necessary to obtain one hundred occurrences of pronouns was taken. He then applied the algorithm to each pronoun and counted the number of times it worked.[7] He reports (1976:25) that the algorithm worked 88 percent of the time, and 92 percent when augmented with simple selectional constraints. In many cases, the algorithm worked because there was only one available antecedent anyway; in the cases where there was more than one, the algorithm combined with selectional restrictions was correct for 82 percent of the time.

Clearly, the algorithm by itself is inadequate. However Hobbs suggests that it may still be useful, as it is computationally cheap compared to any semantic method of pronoun resolution. Because it is frequently necessary for semantic resolution methods to search for inference chains from reference to referent, time may frequently be saved, suggests Hobbs (1976:38), by using a bidirectional search starting at both the reference and the antecedent proposed by the algorithm, seeing if the two paths meet in the middle.

---

[6]From: *Even cowgirls get the blues*. New York: Bantam, 1977, page 379.

[7]To the best of my knowledge, Hobbs is the only worker in NLU to have ever quantitatively evaluated the efficacy of a language understanding mechanism on unrestricted real-world text in this manner. Clearly, such evaluation is frequently desirable.

3.1.7. Preference semantics

Wilks (1973b, 1975a, 1975b) describes an English to French translation system[8] which uses four levels of pronominal anaphor resolution depending on the type of anaphor and the mechanism needed to resolve it. The lowest level, type "A", uses only knowledge of individual lexeme meanings. For example, in (3-18):

> (3-18) Give the bananas to the monkeys although <u>they</u> are not ripe, because <u>they</u> are very hungry.

each *they* is interpreted correctly using the knowledge that monkeys, being animate, are likely to be hungry, and bananas, being a fruit, are likely to be (not) ripe. The system uses "fuzzy matching" to make such judgements; while it chooses the most likely match, future context or information may cause the decision to be reversed. The key to Wilks's system is very general rules which specify PREFERRED choices but don't require an irreversible commitment in case the present situation should turn out to be an exception to the rule.

If word meaning fails to find a unique referent for the pronoun, inference methods for type "B" anaphors — those that need analytic inference — or type "C" anaphors — those that require inference using real-world knowledge beyond simple word meanings — are brought in. These methods extract all case relationships from a template representation of the text and attempt to construct the shortest possible inference chain, not using real-world knowledge unless necessary.

If the anaphor is still unresolved after all this, "focus of attention" rules attempt to find the topic of the sentence to use as the referent.

Wilks's system of rules exhibiting undogmatic preferences, as well as his stratification of resolution requirements, is intuitively appealing, and appears the most promising of the approaches we have looked at; it could well be applied to forms of anaphora other than pronouns. My major disagreement is with Wilks's relegation of (rudimentary) discourse considerations to use only in last desperate attempts. I will show in the next chapter that they need to play a more important role.

3.1.8. Summary

We have seen six basic traditional approaches to anaphora and coreferentiality:

1 a few token heuristics;
2 more sophisticated heuristics with a semantic base;
3 a case-based grammar to give the heuristics extra power, using word meanings as well;

---

[8]For an unbiased description of Wilks's system, see Browse (1976).

4 lots and lots of undirected inference;

5 dumb parse-tree searching, with semantic operations to keep out of trouble;

6 a scheme of flexible preference semantics with word meanings and inference.

In the next section, we will evaluate in greater detail these and other approaches.

> *The Hodja was walking home when a man came up behind him and gave him a thump on the head. When the Hodja turned round, the man began to apologize, saying that he had taken him for a friend of his. The Hodja, however, was very angry at this assault upon his dignity, and dragged the man off to the court. It happened, however, that his assailant was a close friend of the cadi [magistrate], and after listening to the two parties in the dispute, the cadi said to his friend:*
>
> *"You are in the wrong. You shall pay the Hodja a farthing damages."*
>
> *His friend said that he had not that amount of money on him, and went off, saying he would get it.*
>
> *Hodja waited and waited, and still the man did not return. When an hour had passed, the Hodja got up and gave the cadi a mighty thump on the back of his head.*
>
> *"I can wait no longer", he said. "When he comes, the farthing is yours."[9]*

3.2. Abstraction of traditional approaches

Before continuing on to the discourse-oriented approaches to anaphora in the next two chapters, I would like to stand back and review the position so far.

It is a characteristic of research in NLU that, as in many new and smallish fields, the best way to describe an approach is to give the name of the person with whom it is generally associated. This is reflected in the organization of both section 3.1 and Chapter 5. However, in this section I would like to categorize approaches, divorcing them from people's names, and to formalize what we have seen so far.

---

[9]From: Charles Downing (reteller). *Tales of the Hodja*. Oxford University Press, 1964, page 10. This excerpt is recommended for anaphor resolvers not only as a useful moral lesson, but also as a good test of skill and ruggedness.

3.2.1. A formalization of the problem

David Klappholz and Abe Lockman (1975) (hereafter K&L), who were perhaps the first in NLU to even consider the problem of reference as a whole, sketch out the basics of a reference resolver. They see it as necessarily based upon and operating upon representations of meaning, a set of world knowledge and a memory of the FOCUS derived from each past sentence, including noun phrases, verb phrases, and events.[10] One then matches up anaphors with previous noun phrases and other constituents, and uses semantics to see what is a reasonable match and what isn't, hoping to avoid a combinatorial explosion with the aid of the world knowledge.

Specifically, K&L envisaged three focus sets — for noun-objects, events and time.[11] As each sentence comes in, a meaning representation is formed for it; then the focus sets are updated by adding entities from the new sentence, and discarding those from the nth previous sentence, which are now deemed too far back to be referred to. (K&L do not hazard any guess at what a good value for n is.) A hypothesis set of all triples (N_1, N_2, r) is generated, where N_1 is a reference needing resolution, N_2 is an entity in focus and r is a possible reference relation (see section 2.4.2). A judgement mechanism then tries to winnow the hypotheses with inference, semantics and knowledge, until a consistent set is left.

This method is, of course, what Winograd and Woods (see sections 3.1.2 and 3.1.3) were trying to approximate. However, in their formalization of the problem K&L are aiming for higher things, namely a solution for the general problem of definite reference, from which an anaphor resolver will fall out as an immediate corollary. I believe their model however still represents less than the minimum equipment for a successful solution to the problem. For example (as K&L themselves point out) their model cannot handle examples like (2-106)[12] where determining the reference relationship requires inference. Further, as we shall soon see, the model of focus as a simple shift register is overly simplistic.[13]

---

[10]In general, we will mean by the FOCUS of a point in text all concepts and entities from the preceding text that are referable at that point. As should soon be clear, focus is just what we have been calling "consciousness".

[11]In Hirst (1976b), I proposed that their model really requires three other focus sets — locative, verbal and actional — for the resolution of locative, pro-verbial and proactional anaphors, respectively.

[12](2-106) "It's nice having dinner with candles, but there's something funny about the two we've got tonight", Carol said. "They were the same length when you first lit them. Look at them now."
   John chuckled. "The girl did say one would burn for four hours and the other for five", he replied . . .

[13]K&L have since developed their model to eliminate some of these problems, and we will see their later work in section 5.5. My reason for presenting their earlier work here is that it serves as a useful conceptual scaffold from which to build both our review of traditional anaphora resolution methods and our exposition of modern methods.

3.2.2. Syntax methods

Linguists have found many syntactic constraints on pronominalization in sentence generation. These can be used to eliminate otherwise acceptable antecedents in resolution fairly easily. We will look at a couple of examples:[14]

The most obvious constraint is REFLEXIVIZATION. Consider:

(3-19) Nadia says that Sue is knitting a sweater for her.

Her is Nadia or, in the right context, some other female, but cannot be Sue, as English syntax requires the reflexive *herself* to be used if Sue is the intended referent. In general an anaphoric NP is coreferential with the subject NP of the same simple sentence if and only if the anaphor is reflexive.

Another constraint prohibits a pronoun in a main clause referring to an NP in a subsequent subordinate clause:

(3-20) Because Ross slept in, he was late for work.

(3-21) Because he slept in, Ross was late for work.

(3-22) Ross was late for work because he slept in.

(3-23) He was late for work because Ross slept in.

In the first three sentences, *he* and *Ross* can be coreferential. In (3-23), however, *he* cannot be Ross because of the above constraint, and either *he* is someone in the wider context of the sentence or the text is ill-formed.

We have already seen that syntax-based methods by themselves are not enough. However, syntax-oriented methods may still play a role in anaphora resolution, as we saw in section 3.1.6.

> *The fool hath said in his heart,*
> *There is no God.*
> *— David*[15]

3.2.3. The heuristic approach

This is where prejudices start showing. Many AI workers, myself included, adhere to the maxim "One good theory is worth a thousand heuristics". People like Yorick Wilks (1971, 1973a, 1973b, 1975c) would disagree, arguing that language by its very nature — its lack of a sharp boundary — does not always allow (or perhaps NEVER allows) the formation of "100%-correct" theories; language understanding cannot be an exact science, and therefore heuristics will always be needed to plug the gaps. If the heuristic approach has failed so

---

[14]See Langacker (1969) and Ross (1969) for more syntactic restrictions on pronominalization.

[15]*Psalms* 14:1.

far, so this viewpoint says, then we just haven't found the right heuristics.[16]

While not totally rejecting Wilks's arguments,[17] I believe that the search for a good theory on anaphora resolution should not yet be terminated and labelled a failure. Gathering heuristics may suffice for the construction of a particular practical system, such as LSNLIS, but the aim of present work is to find more general principles. (Chapter 5 describes several theoretical approaches to the problem.)

This does not mean that we have no time for heuristics. The essence of our quest is COMPLETENESS. Thus, a taxonomy of anaphors or coreferences, together with an algorithm which will recognize each and apply a heuristic to resolve it, would be acceptable if it could be shown to handle every case the English language has to offer. And indeed, if we were to develop the heuristic approach, this would be our goal.[18]

However, our prospects for reaching this goal appear dismal. Consider first the problem of a taxonomy of anaphors, coreferences and definite references. Halliday and Hasan (1976), in attempting to classify different usages in their study of cohesion in English, identify 26 distinct types which can function in 29 distinct ways. (Compare my loose and informal classification in section 2.3.15.) While it is possible that some of their categories can be combined in a taxonomy useful for computational understanding of text, it is equally likely that as many, if not more, of their categories will need further subdivision. There is, moreover, no way yet of ensuring completeness in such a taxonomy, nor of ensuring that a heuristic will work properly on all applicable cases.

Also, there is the problem of semantics again. Rules which will allow the resolution of anaphors like those of the following examples will require either a further fragmentation of the taxonomy, or a fragmentation within the heuristic for each category:

> (3-24) When Sue went to Nadia's home for dinner, she served sukiyaki au gratin.
>
> (3-25) When Sue went to Nadia's home for dinner, she ate sukiyaki au gratin.

(These examples will be referred to collectively below as the 'sukiyaki' examples.) Here she, superficially ambiguous, means Nadia in (3-24) and Sue in (3-25).

Thus, a heuristic approach will essentially degenerate into a demon-like system (Charniak 1972), in which each heuristic is just a demon watching out for its own special case. Although this is theoretically fine, the shortcomings of such systems are well-known (Charniak 1976).

---

[16] For a discussion of Wilks's arguments in detail, see Hirst (1976a).

[17] I confess that when in a slough of despond I sometimes fear he may be right.

[18] One attempt at the heuristic approach was made by Baranofsky (1970), who described such a taxonomy with appropriate algorithms. However, her heuristics made no attempt to be complete, but rather to cover a wide range with as few cases as possible. I have been unable to determine whether the heuristics were ever implemented in a computer program.

3.2.3 The heuristic approach

All this is not to do away with heuristics entirely. As Wilks points out, we may be forced to use them to plug up holes in any theory, and, moreover, any theory may contain one or more layers of heuristics.[19]

3.2.4. The case grammar approach

Case "grammars" (Fillmore 1968, 1977), with their wide theoretical base, are able to resolve many anaphors in a way that is perhaps more simple and elegant than heuristics. The extra information provided by cases is often sufficient to easily pair reference with referent, given the meaning of the words involved.

For example, this approach is able to handle differences in the meaning of a word or anaphor in context. Compare (3-26) and (3-27):

(3-26) Ross asked Daryel to hold <u>his</u> books for a minute.

(3-27) Ross asked Daryel to hold <u>his</u> breath for a minute.

In the first sentence, *his* refers to Ross, the default referent,[20] and in the second, it refers to Daryel. Further, in each sentence, *hold* has a different meaning — *support* and *retain* respectively[21] — and handling the difference would be difficult for many systems. A case-driven parser, such as Taylor's (1975) (see section 3.1.5), would have a dictionary entry for each meaning of *hold*. In this example, *breath* could only pass the tests associated with the case-frame for one meaning, while *books* could only pass the tests for the other. Hence the correct meaning would be chosen. It is then possible to resolve the anaphors. In (3-26), there is nothing to contraindicate the assignment of the default. In (3-27), the system could determine that since the *retain* sense of *hold* was chosen, *his* must refer to Daryel. Taylor's parser does not have this resolution capability, but to program it would be fairly straightforward, if a default finder could be given.

Case-based systems also have an advantage in the resolution of situational anaphors. Compare (2-40)[22] with (3-28):

---

[19] You may have noticed that most of my arguments in this section depend on precisely what I mean by a "heuristic", and that I have placed it somewhere on a continuum between "theory" and "demon". While this is not the place to discuss this matter in detail, I am using the word to mean one of a set of essentially uncoordinated rules of thumb which together suffice to provide a method of achieving an end under a variety of conditions.

[20] Some idiolects appear not to accept this default, and see the anaphor as ambiguous.

[21] That these two uses of *hold* are not the same is demonstrated by the following examples:
  (i) Daryel held his books and his briefcase.
  (ii) ?Daryel held his books and his breath.

[22] (2-40) The president was shot while riding in a motorcade down a major Dallas boulevard today; <u>it</u> caused a panic on Wall Street.

(3-28) The president was shot while riding in a motorcade down a major Dallas boulevard today; <u>it</u> was crowded with spectators at the time.

A general heuristic system would have trouble detecting the difference between the *it* in each case. A case grammar approach can use the properties of the verb forms *to be crowded* and *to cause* to recognize that in (2-40) the referent may be situational. To determine exactly what situation is being referred to, though, some UNDERSTANDING of sentences will be needed. This problem doesn't arise in this particular example, since there is only one previous situation that can be referenced prosententially. But as we have seen, whole paragraphs and chapters can be prosententially referenced, and deciding which previous sentence or group of sentences is intended is a task which requires use of meaning.

The case approach would not be sufficient to resolve our 'sukiyaki' examples. Recall (3-24).[23] The parser would look for a referent for *she* with such conditions as MUST-BE HUMAN, MUST-BE FEMALE and SHOULD-BE HOST. But how is it to know that Nadia, and not Sue, is the item to be preferred as a HOST? Humans know this from the location of the event taking place. However, a case-driven parser does not have this knowledge, expressed in the subordinate clause at the start of the sentence, available to it. To get this information, an inferencing mechanism is needed to determine from the verb *went* that the serving took place at, or on the way to,[24] Nadia's home, and to infer that therefore Nadia is probably the host. Such an inferencer will also need to use a database of information from previous sentences, as not all the knowledge necessary for resolution need be given in the one sentence at hand. (For example, in this case the sentence may be broken into two simple sentences.) This database must contain semantic information — meanings of, and inferences from, past sentences; that is, sentences must be, in some sense, UNDERSTOOD.[25] Thus we see once more that parsing with anaphor resolution cannot take place without understanding.

Now consider (3-25).[26] Here, a case approach has even less information — only MUST-BE ANIMATE and MUST-BE FEMALE — and no basis for choosing between Sue and Nadia as the subject of the main clause. The way we know that it is Sue is that she is the topic of the preceding subordinate clause and, in the absence of any indication to the contrary, the topic remains unchanged. Notice that this rule is neither syntactic nor semantic but pragmatic — a convention of conversation and writing. Apart from this, there is no other way of determining that Sue, and not Nadia, is the sukiyaki consumer in question.

---

[23] (3-24) When Sue went to Nadia's home for dinner, <u>she</u> served sukiyaki au gratin.

[24] Sentence (i) shows that we cannot conclude from the subordinate clause that the location of the action expressed in subsequent verbs necessarily takes place at Nadia's home:

  (i) When Sue went to Nadia's home for dinner, she caught the wrong bus and arrived an hour late.

[25] The database will also need common-sense real-world knowledge.

[26] (3-25) When Sue went to Nadia's home for dinner, <u>she</u> ate sukiyaki au gratin.

3.2.4 The case grammar approach

Another use of cases is in METAPHOR resolution for anaphor resolution. A system which uses a network of cases in conjunction with a network of concept associations to resolve metaphoric uses of words has been constructed by Roger Browse (1977, 1978). For example, it can understand that in:

(3-29) Ross drank the bottle.

what was drunk was actually the contents of the bottle. This is determined from the knowledge that bottles contain fluid, and *drink* requires a fluid object. Such metaphor resolution can be necessary in anaphor resolution, especially where the anaphor is metaphoric but its antecedent isn't, or vice versa. For example:

(3-30) Ross picked up the bottle and drank it.
(3-31) Ross drank the bottle and threw it away.

We can conclude from this discussion that a case-base is not enough, but a maintenance of focus (possibly by means of heuristics) and an understanding of what is being parsed are essential. We have also seen that cases can aid resolution of metaphoric anaphors and anaphoric metaphors.

How could such a case system resolve paraphrase coreferences and definite reference? Clearly, case information alone is inadequate, and will need assistance from some other method. Nevertheless, we see that a case "grammar" may well serve as a firm base for anaphora resolution.

3.2.5. Analysis by synthesis

Transformational grammarians have spent considerable time pondering the problem of where pronouns and other surface proforms come from, and have produced a number of theories which I will not attempt to discuss here. This leads to the possibility of anaphora resolution through analysis by synthesis, where we start out with an hypothesized deep structure which is generated by intelligent (heuristic?) guesswork, and apply transformational rules to it until we either get the required surface or fail.

What this involves is a parser, such as the ATN parser of Woods (1970), to provide a deep structure with anaphors intact. Then each anaphor is replaced by a hypothesis as to its referent, and transformations are applied to see if the same surface is generated. If so, the hypotheses are accepted; otherwise new ones are tried. The hypotheses are presumably selected by a heuristic search.

There are many problems with this method. First, the generation of a surface sentence is a nondeterministic process which may take a long time, especially if exhaustive proof of failure is needed; a large number of combinations of hypotheses may compound this further. Second, this approach does not take into account meanings of sentences, let alone the context of whole paragraphs

or world knowledge. For example, in (3-32):

(3-32) Sue visited Nadia for dinner because <u>she</u> invited <u>her</u>.

both the hypotheses *she = Sue, her = Nadia* and *she = Nadia, her = Sue* could be validated by this method and without recourse to world knowledge there is no way of deciding which is correct. Third, the method cannot handle intersentential anaphora. We must conclude that analysis by synthesis is not promising.

3.2.6. Resolving anaphors by inference

If we are to bring both world knowledge and word meaning to bear in anaphora resolution, then some inferencing mechanism which operates in this domain is needed. Possible paradigms for this include Rieger's Conceptual Memory (1975) (see 3.1.4) and Wilks's preference semantics (1973b, 1975a, 1975b) (see section 3.1.7).

Although conceptual dependency, which Conceptual Memory uses, is not without its problems (Davidson 1976), it may be possible to extend it for use in anaphor resolution. This would require giving it a linguistic interface such that reasoning which involves world knowledge, sentence semantics and the surface structure can be performed together — clearly pure inference, as in Conceptual Memory, is not enough. An effective method for representing and deploying world knowledge will also be needed. A system using FRAMES (Minsky 1975), or SCRIPTS (Schank and Abelson 1975, 1977) (which are essentially a subset of frames), appears promising. Frames allow the use of world knowledge to develop EXPECTATIONS about an input, and to interpret it in light of these. For instance, in the 'sukiyaki' examples, the mention of Sue visiting Nadia's home should invoke a VISITING frame, in which the expectation that Nadia might serve Sue food would be generated, after which the resolution of the anaphor is a matter of easy inference.

In Wilks's system inference is more controlled than in Conceptual Memory; whereas the latter searches for as many inferences to make as it can without regard to their possible use,[27] the former tries to find the shortest possible inference chain to achieve its goal. Although Wilks's system does not use the concept of expectations, its use of preferred situations can achieve much the same ends. In the 'sukiyaki' examples, the host would be the preferred server.

---

[27]Rieger has since developed a more controlled approach to inference generation (Rieger 1978).

3.2.7. Summary and discussion

I have discussed in this section five different approaches to anaphor resolution. They are:

1 syntactic methods — which are clearly insufficient;

2 heuristics — which we decided may be necessary, though we would like to minimize their inelegant presence, preferring as much theory as possible;

3 case grammars — which we saw to be elegant and powerful, but not powerful enough by themselves to do all we would like done;

4 analysis by synthesis — which looks like a dead loss; and

5 inference — which seems to be an absolute necessity to use world knowledge, but which must be heavily controlled to prevent unnecessary explosion.

From this it seems that an anaphor resolver will need just about everything it can lay its hands on — case knowledge, inference, world knowledge, and word meaning to begin with, not to mention the mechanisms for focus determination, discourse analysis, etc that I will discuss in subsequent chapters, and perhaps some of the finer points of surface syntax too.[28]

---

[28]That a boots-and-all approach is necessary should perhaps have been clear from the earliest attempts in this area because of the very nature of language. For natural language was designed (if I may be so bold as to suggest a high order of teleology in its evolution) for communication between human beings, and it follows that no part of language is beyond the limits of competence of the normal human mind. And it is not unreasonable to expect, a fortiori, that no part is far behind the limits of competence either, for if it were, either it could not meet the need for a high degree of complexity in our communication, or else language use would be a tediously simplistic task requiring long texts to communicate short facts.

Consider our own problem, anaphora. Imagine what language would be like if we did not have this device to shorten repeated references to the same thing, and to aid perception of discourse cohesion. Clearly, anaphora is a highly desirable component of language. It is hardly surprising then that language should take advantage of all our intellectual abilities to anaphorize whenever it is intellectually possible for a listener to resolve it. Hence, any complete NLU system will need just about the full set of human intellectual abilities to succeed. (See also Rieger (1975:268).)