AN EXPLANATION FOR MINIMAL ATTACHMENT AND RIGHT ASSOCIATION

Robert W. Milne
Air Force Institute of Technology
AFIT/ENG
Wright-Patterson AFB, OH 45433

ABSTRACT

This paper describes the principles of Right Association and Minimal Attachment and explains how the Sausage Machine and ATN decribe these principles. It is then shown that these two models cannot explain these principles. It is then shown that a production system grammar can both describe these principles as well as suggest why they must be true.

A. The Sausage Machine

Lyn Frazier and Janet Fodor (FF) in [1] have proposed a two stage model of the Human Sentence Parsing Mechanism (HSPM), called the Sausage Machine (SM).

FF proposed that the syntactic analysis of sentences by hearers or readers is performed in two stages. The first stage combines words in phrasal nodes as they are received. They call this the "Preliminary Phrase Packager" (PPP) or the "Sausage Machine". The second stage combines these phrases into sentences. This stage is called the "Sentence Structure Supervisor" (SSS).

The main principles of Frazier and Fodor are these:

1) Right Association... an ambiguous constituent should be "attached into the phrase marker as a right sister to existing constituents and as low in the tree as possible" [6 p.2111] [1 p. 294]

2) Minimal Attachment...an ambiguous item "is to be attached into the phrase marker with the fewest possible number of non-terminal nodes linking it with the nodes already present" [1 p. 320]

---

*This paper describes work done in the Department of Artificial Intelligence and School of Epistemics, Edinburgh University, Scotland

Right Association suggests that "terminal symbols optimally associate to the lowest non-terminal node." This predicts the preferred interpretation of:

[1] Tom said that Bill had taken the cleaning out yesterday.
[2] Joe called the friend who smashed his new car up.
[3] John read the note, the memo and the letter to Mary.
[4] The girl took the job that was attractive. (from [1 p. 297])

In each of these sentences, the preference is to attach the final modifier to the lowest right node, just as their principle predicts. This principle also predicts the difficulty in the following sentences.

[5] Joe looked the friend who had smashed his new car up.
[6] John read the note, the memo and the newspaper to Mary.
[7] The girl applied for the job that was attractive.

Their second main point is "Minimal Attachment" (MA). This says "Each lexical node (or other node) is to be attached into the Phrase marker with the fewest possible number of non-terminal nodes linking it with the nodes which are already present. [1 p. 320] This principle accounts for the preferred attachment of (for Susan) to the VP in:

[8] John bought the book for Susan.

They suggested that this accounted for the preference for the conjunctive analysis of NP NP in center embedded sentences, the preference for the first clause to be a main one, as in most garden paths, and the preference [5] for "that" as a complementiser rather than a relative clause when after a NP. It even predicts the use of "that" as a determiner over the "comp" usage.

B. The ATN Response

Wanner has replied to their claims and defended the ATN as a model of human sentence parsing in [6].

Wanner claims that the SM does not model Minimal Attachment and Right Association in a principled and independent way. In this paper he demonstrated that the ATN can.

In accounting for this data, Wanner first presented the background of the ATN and presented these possible arc types:

WORD arc - analyse specific words such as "that" or "to"
CAT arc - analyse grammatical categories such as Noun and Verb
SEEK arc - analyse whole phrases such as NP,VP or S
SEND arc - terminate a network (node)
JUMP arc - express optionality [6 p. 216]

He then stated the two principles in terms of these arcs:

Right Association: Schedule all SEND and JUMP arcs after every other type of arc.

Minimal Attachment: Schedule all CAT and WORD arcs before all SEEK arcs.

He presented an argument that shows that this characterisation of the arcs is correct and FF agreed basically with this re-formulation. [2]

Wanner has shown that the ATN can describe the strategies of MA and RA, but cannot explain why these strategies are present. So even though the ATN can show sufficient description, it cannot show necessity. Wanner even wonders this when he asked: "Why does the parser employ these strategies as opposed to others?" [6 p. 233] He then admitted that no clear answer was available to this question for an ATN. Wanner wondered why these are true and had trouble suggesting an adequate explanation of why they are necessary.

C. The Production System Reply

Milne [3] presents a deterministic parser, ROBIE, modeled after PARISIFAL [4]. Both parsers are deterministic and use limited lookahead as well as a production system grammar. In the next paragraphs, I will show that not only can ROBIE describe the principles above in the same way, but it can explain why they must be true. We will look at these principles one at a time.

In some production systems, the following principle gives the order in which to test each rule: "The most constrained pattern is tried first". For if the rules were tried in the opposite order, the more constrained rules would never be matched. This provides an order for rules of unequal constraint. The following rules will be tried in the following order:

[so] [that] ->   <action1>
[to] [tenseless] ->   <action2>
[that] ->   <action3>
[noun] [noun] ->   <action4>
[noun,npl] ->   <action5>
[verb] ->   <action6>
[t] ->   <action7>

A pattern with one word is more constrained than a pattern with two features, since there is only one lexical item that can match the first, but several lexical items that may have, say "tenseless". A rule with no pattern [t] will always be the tried last. This is necessary to handle many ambiguity issues. For, if the rules were tried in the opposite order, the more constrained rules would never be matched. I emphasize, this same principle says that all default rules (rules with no pattern) will have lower priority than any other rule.

1. Right Association

ROBIE does not have the types of arcs that were listed for the ATN earlier, but the rules can be divided into several roughly similar groups. The equivalent of the SEND and JUMP arcs would be the default rules in a packet (Sub set of the grammar). If something is optional, typically a packet has a rule to handle the marked case and a default rule to handle the unmarked case, that is the default rule has no pattern. All equivalents of the SEND and JUMP arcs will have no pattern in the current parser. Hence, according to our above ordering, these rules will be tried last. Thus Wanner's explanation of Right Association is a necesary result of ROBIE's design.

2. Minimal Attachment

The deterministic parser has no SEEK arcs. A grammar rule in ROBIE with the pattern [np], does not create a NP. Instead, this pattern will match only if a NP node has already been started. But in the ATN, the arc with the NP on it will cause a push to the NP subnetwork and try to build a NP. Ordering this SEEK arc is the problem under discussion here.

On the subject of no SEEK arcs, Marcus states:

"The pattern that triggers on a specific constituent, says a NP or an S, does not initiate parsing of a constituent of that sort. Instead, the pattern will only trigger if a constituent of that sort is already in the specified buffer." [4 p.22]

If a pattern has the feature NP, this does not make ROBIE try to parse an NP. Instead the pattern will match only if a node with that feature has already been built. This can be contrasted with the SEEK (or PUSH) arc of the ATN. The SEEK arc tries to build a node of the type that was specified on it. SEEK arcs are like recursive subroutine calls.

Because ROBIE does not have SEEK arcs, the problems of ordering them are not relevant. The CAT and WORD arcs will be scheduled first as Wanner has shown necessary.

MA as characterised by Wanner states that essentially the parser should be data driven and should reflect the incoming words. Another way to understand the principle of Minimal Attachment, is that the word should be used locally if it fits.

Since ROBIE has no access to the Active Node Stack, except for the active packets, then it is unable to see if the word could be used higher up. If the word, could be attached to the lower node, then the grammar rules must be written to handle it there. If these rules are there, then the optional use will be grabbed and this will behave exactly as Minimal Attachment.

D. Summary

In this paper we have seen that while the principles of MA and RA can be described by arc ordering, why this order must be used is not obvious. We then discussed that the only possible order of a production system grammar provides the same arc ordering necessary to describe MA and RA. This suggests that RA and MA are merely side effects of a grammar design and not a specific parsing mechanism.

Bibliography

[1] Fodor, J. and Frazier, L. "The Sausage Machine: A New Two-Stage Parsing Model", in Cognition 6. (1970)

[2] Fodor, J. and Frazier, L, "Is the Human Sentence Parsing Mechanism an ATN?", Cognition 8. (1980)

[3] Milne, R. [forthcoming] "Resolving Lexical Ambiguity in a Deterministic Parser", PhD Thesis, Dept of AI, Edinburgh University

[4] Marcus, M. A Theory of Syntactic Recognition for Natural Language, MIT Press. (1980)

[5] Wanner, E., Kaplan, J., Shiner "Garden Paths in Relative Clauses", unpublished paper, Harvard University. (1975)

[6] Wanner, E. "The ATN and The Sausage Machine: Which one is Baloney", Cognition 8. (1980)