Proceedings of the Annual Meeting of the Cognitive Science Society

Title

A Parallel Constraint Satisfaction and Spreading Activation Model for Resolving Syntactic Ambiguity

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 12(0)

Author

Stevenson, Suzanne

Publication Date

1990

A Parallel Constraint Satisfaction and Spreading Activation Model for Resolving Syntactic Ambiguity

Suzanne Stevenson
Department of Computer Science
University of Maryland, College Park

Abstract

This paper describes a computational architecture whose emergent properties yield an explanatory theory of human structural disambiguation in syntactic processing. Linguistic and computational factors conspire to dictate a particular integration of symbolic and connectionist approaches, producing a principled cognitive model of the processing of structural ambiguities. The model is a hybrid massively parallel architecture, using symbolic features and constraints to encode structural alternatives, and numeric spreading activation to capture structural preferences. The model provides a unifying explanation of a range of serial and parallel behaviors observed in the processing of structural alternatives. Furthermore, the inherent properties of active symbolic and numeric information correspond to general cognitive mechanisms which subsume a number of proposed structural preference strategies.

1 Introduction

It has been repeatedly demonstrated that people have little trouble in processing structurally ambiguous sentences; moreover, they yield consistent structural preferences in the face of ambiguity. Yet theories of human structural preferences have progressed little beyond the stage of unrelated descriptions of each piece of the psychological data. The research described here aims to shed light on the cognitive principles used in structural disambiguation by exploring the computational mechanisms which underlie them. The goal is not to create a parser in which human structural preferences are built in, but to design a parsing architecture whose basic properties *predict* those preferences.

A predictive theory of structural disambiguation emerges from the active, distributed nature of the computational model described here. The model is a hybrid massively parallel architecture, combining symbolic constraint satisfaction and numeric spreading activation. Symbolic features and constraints based on Chomsky's (1981, 1986) Government-Binding theory (GB) capture multiple structural alternatives in parallel in a linguistically motivated way. Numeric spreading activation guides the choice between these structural alternatives by encoding and integrating the degree of featural compatibility, the recency of activation, and the strength of lexical preference for each possible attachment. The preference for a particular structural attachment is thus uniformly determined by an inherent mechanism of the architecture.

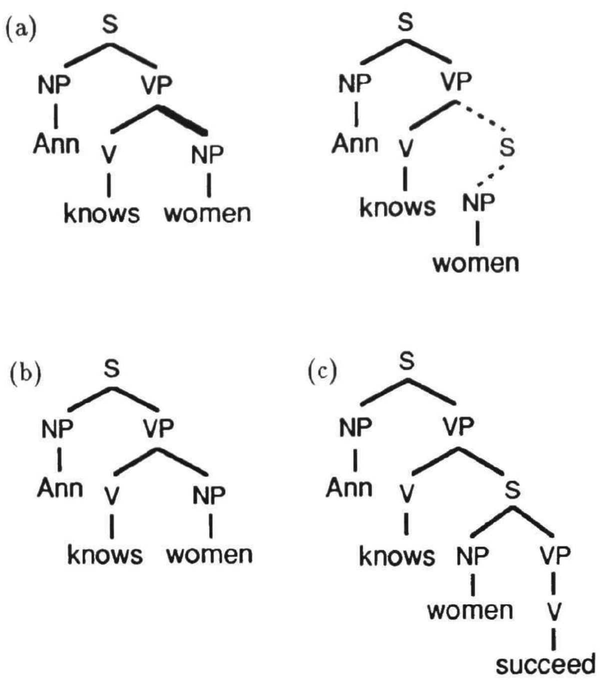
The remainder of this paper discusses the model and its consequences in more detail. Section 2 describes the key psycholinguistic issues which must be addressed by an explanatory theory of structural disambiguation. Section 3 presents the details of the hybrid architecture and describes the properties from which its structural disambiguation behavior emerges. In Section 4, the explanatory power of the mechanism is demonstrated. Section 5 discusses some related work on structural disambiguation, and Section 6 concludes the paper with a summary of the contributions of this research.

2 The Psycholinguistic Data

A structural ambiguity gives rise to multiple attachment possibilities for a syntactic phrase. Any model of structural disambiguation must address the two issues of how to process the valid structural alternatives for the phrase, and how to capture the structural preference factors which choose between them. This section describes some of the psycholinguistic observations related to these issues.

Serialism versus parallelism

One of the first issues that must be addressed in building a predictive model of structural preferences is the degree to which structural alternatives are processed in parallel. This is, in fact, a major open question in psycholinguistic research. Clear evidence for a serial mechanism comes from experiments which demonstrate consistent strong preferences for one resolution over another of a temporary ambiguity (Frazier, 1978). For example, in the sentence beginning:



(a) ▬ preferred ------ non-preferred

(b) preferred resolution of the ambiguity.

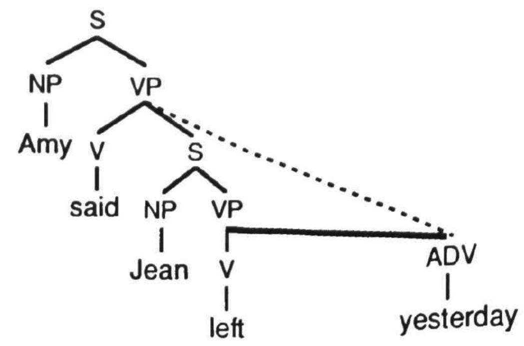
(c) non-preferred resolution of the ambiguity.

Figure 1: Attachment possibilities for *women*.

Ann knows $[_{\text{NP}}$ women$]$ …,

the attachment of the NP is temporarily ambiguous, as shown in Figure 1. The sentence may end after *photo* or continue "…succeed," each case resolving the ambiguity in a different way. The consistent preference for the first of these resolutions of the NP attachment indicates that the parser chooses one of the possible structural alternatives and pursues it serially.

However, equally convincing evidence for a parallel architecture emerges from experiments which reveal the availability of multiple structures in on-line processing of these temporary ambiguities (Gorrell, 1987). That is, experiments indicate that in the above example, both the NP–to–VP and the NP–to–S attachment possibilities are maintained in parallel. A major contribution of the model proposed here is that it naturally accounts for these apparently contradictory results, while other models have failed to do so. Exploiting the interesting interaction of serial and parallel qualities in fact leads to the model's ability to provide an architectural explanation for the range of serial and parallel behaviors attested in a number of psycholinguistic experiments. Section 3 describes the hybrid model which turns this tension between serial and parallel processing to its advantage.



▬ preferred ------ non-preferred

Figure 2: Attachment possibilities for *yesterday*.

Structural preference accounts

Numerous principles have been proposed to account for the consistent structural preferences displayed by the human parser in the face of syntactic ambiguity. For example, Minimal Attachment asserts that when the parser has a choice between two or more ways of attaching the current phrase into the parse tree, it will pick the one which requires the creation of the fewest number of new nodes (Frazier, 1978). The preference in Figure 1 for the NP to attach as the object of the verb is a clear case of Minimal Attachment: the parser prefers to attach the NP directly to the VP, rather than creating the S node to serve as its attachment site.

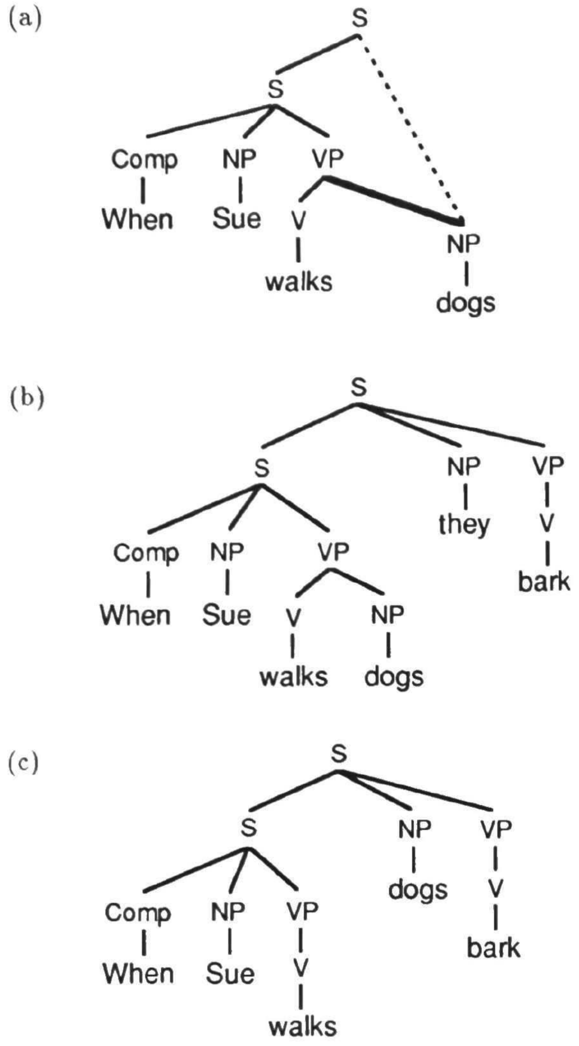
Another principle, Late Closure, states that the parser will prefer to keep a constituent open (that is, available to attach into) as long as possible, entailing that people will prefer to attach a phrase into the most recent open constituent (Frazier & Rayner, 1982). Late Closure accounts for the preferences indicated in Figures 2 and 3. In each case, the non-preferred attachment would require first closing off the open subordinate verb phrase.

Theories of lexical preferences claim that verbs have varying strengths of expectation for their possible arguments. For example, in these sentences taken from Ford, Bresnan, & Kaplan (1982):

Joe included $[_{\text{NP}}$the package $[_{\text{PP}}$for Susan$]]$.
Joe carried $[_{\text{NP}}$the package$]$ $[_{\text{PP}}$for Susan$]$.

the PP attachment preferences indicated by the given bracketings result from a difference in how strongly the verbs *include* and *carry* expect a PP argument.

These are just a few of the many such accounts of the factors involved in structural disambiguation (for example, Kimball, 1973; Frazier & Fodor, 1978; Nicol, 1988; McRoy & Hirst, 1989). Although valuable as concise descriptions of a wealth of psycholinguistic phenomena, these accounts do not explain *why* the parser



(a) —— preferred ------ non-preferred

(b) preferred resolution of the ambiguity.

(c) non-preferred resolution of the ambiguity.

Figure 3: Attachment possibilities for *dogs*.

follows these particular patterns of structural preferences and not others. On closer examination, many of these psycholinguistic principles can be shown to be specific statements of the results of more general cognitive processes: the impetus to immediately structure incoming material, the decrease in salience of a structure over time, and the increase in salience given higher frequency or priming. In the model described here, these processes are precisely mirrored in its inherent qualities of active distributed computation, decay of activation, frequency-encoding weights, and activation of expectations. The properties of its active distributed computation are in turn strongly influenced by the model's integration of serial and parallel processing. Section 4 presents in detail how these emergent properties of the architecture predict the observed pattern of preferences.

3 The Hybrid Architecture

The rise over the past decade of the connectionist approach to cognitive modeling has generated much debate over the relative merits of serial symbolic processing models and massively parallel architectures restricted to numeric spreading activation. The debate has sparked an interest in so-called "hybrid" models which attempt to exploit the desired properties of each approach, while avoiding their respective pitfalls. A precise formulation of the components of the structural disambiguation process has motivated the design of a hybrid model in the research presented here.

Structural disambiguation involves two distinct parsing processes: identifying the allowable attachments for a phrase, and choosing between them. The first process involves the linguistic *competence* of the parser; it uses grammatical knowledge to select the valid structural alternatives. The second process is a matter of linguistic *performance*; extragrammatical factors are taken into account in determining the structural preferences. The motivation for a hybrid approach to a computational explanation of structural disambiguation arises from the necessity of capturing within a single model the properties relevant to both of these processes. Traditional serial symbolic processing models have been good at encoding discrete competence knowledge, while connectionist models are quite successful at integrating the multitude of factors affecting performance.

The question, of course, is how to combine these divergent approaches in a principled way. This can be achieved more naturally than might be expected. Properties of competence and performance themselves each converge on some type of massively parallel architecture. On the competence side, a recent trend in linguistic theory has been away from unwieldy, construction-specific rule-based systems toward a so-called "principles and parameters" approach. Government-Binding theory (GB), founded on this approach, is a constraint-based theory in which the validity of syntactic structures is determined by local licensing relations among constituent phrases. An active, distributed architecture lends itself well to the formulation of grammatical knowledge as a set of simultaneous declarative constraints which must be satisfied locally.

On the performance side of the issue, processing structural attachments requires some interesting interaction of serialism and parallelism, as noted in Section 2. Spreading activation through a parallel network inherently combines aspects of serial and parallel processing. Although highly parallel in its simultaneous communication to all neighboring nodes, activation is intrinsically serial in its spread through the space of the network. Furthermore, the massive parallelism of activation must be harnessed through some kind of focusing mechanism, or its sole effect will be network saturation.

Thus the problem domain itself strongly supports a massively parallel architecture combining symbolic constraint satisfaction and numeric spreading activation.

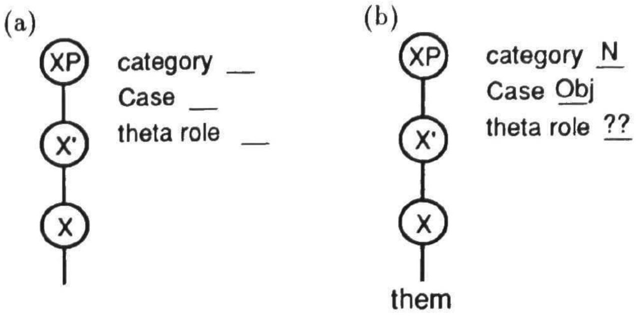
The design of a hybrid architecture must address the issue of how to integrate the seemingly incompatible properties of the symbolic and connectionist processing paradigms. As shown in Table 1, traditional symbolic models typically manipulate symbols serially, building new structure to solve a problem. Connectionist models, on the other hand, compute numeric activation functions in parallel, solving problems by activating the appropriate built-in structures. In the model described here, linguistic and computational principles have converged on a profitable synthesis of these approaches along each of the three relevant dimensions.

First consider the units of information. In the model, symbolic constraint satisfaction naturally encodes GB, while numeric spreading activation acts as a uniform mechanism to capture diverse sources of structural preference information. Many possibilities have previously been explored for effectively integrating symbolic and numeric computation in a cognitive model: Symbolic and numeric computation may operate at different levels of abstraction (Hendler, 1987); they may operate at the same level of abstraction, but independently (Waltz & Pollack, 1985); and activation may constrain the passing of symbolic information (Hendler, 1986). The model here incorporates a new approach, in which symbolic features constrain the spreading of activation. Symbolic constraint satisfaction directly affects the numeric activation of a node, and determines the paths along which activation can spread beyond the node. This is accomplished by using an activation function which depends in part on the *state* of a node, which is a numeric estimation of its degree of constraint satisfaction.

Both linguistic and computational reasons motivate this technique. Symbolic features, which represent linguistic competence, control what is affected by numeric activation, which guides performance. The primacy of symbolic information has a positive computational effect, since it restrains the unwanted spreading of activation. For example, many nodes represent potential syntactic attachments which will be determined to not satisfy the necessary grammatical constraints. If activation could spread across these bad attachments, lending other nodes false support, it would make the determination of structural preferences based on activation much more difficult. Instead, a negative *state*, indicating invalid feature values, forces a node's activation to zero even when it is receiving external stimuli.

Symbolic Processing	Connectionism
symbols	numeric activation
serialism	parallelism
dynamic structure (creation)	fixed structure (recognition)

Table 1: Properties of Symbolic Processing and Connectionist Models.



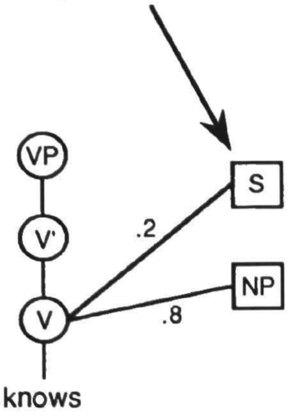
(a) $\overline{\text{X}}$ template and sample features.

(b) template instantiated and initialized by *them*.

Figure 4: The parser's generic syntactic phrases.

Second, the issue of dynamic versus fixed structure must be resolved. Most massively parallel parsers are based on a fixed network of nodes (Cottrell, 1985; Selman & Hirst, 1985); only the model of Waltz & Pollack (1985) has made use of dynamic structure creation. However, their model uses a traditional serial structure-building parser to construct a network corresponding to the parse(s) of the input. The model here strikes a compromise between a totally fixed network structure and the ability to create an arbitrary network structure on the fly. The parser is limited to a single fixed phrase structure template, but it may instantiate this template at will and connect the instances to the input. This "generic" syntactic phrase is then passed initializing features by the associated input. Figure 4 shows the template and a sample instantiation. All logical possibilities of inter-phrase attachments are represented by dynamically allocated attachment nodes; constraint satisfaction rules out those attachments that are invalid.

Once more, this approach is motivated by both linguistic and computational factors. The generic phrasal template is inspired by the lack of general phrase structure rules in GB. $\overline{\text{X}}$ theory, a subsystem of GB, conceives of all phrases as having the same fixed structural shape, with differences in grammatical behavior entailed by features projected from the input. From a computational perspective, this vastly simplifies the structure-building component by restricting it to uniform instantiation of fixed templates. The approach also incorporates a lexically-driven aspect which allows the model to respond to conditions in the input in a straightforward way. For example, a verb can easily determine the weights on its connections to attachment nodes, dynamically taking into account the frequency



Attachment nodes corresponding to the expectations of the verb.

The verb sets the weights on the connections to nodes which represent potential attachments.

Figure 5: Weights capture lexical frequencies.

of its various arguments, as shown in Figure 5.

Finally, the model incorporates elements of both serial and parallel computation. In other massively parallel parsers, the serial aspect of spreading activation is the only constraint on the parallelism of the computation. Here the parallelism is further restricted, by prohibiting top-down precomputation of phrase structure. An $\overline{X}$ node may trigger multiple attachment alternatives, but it cannot cause instantiation of phrase structure based solely on an expectation. For example, in Figure 5 the verb may establish attachment nodes corresponding to its NP and S expectations, but it cannot license the building of an $\overline{X}$ phrase for either of these alternatives. Although quite restricted, the model is still highly parallel in that alternative attachments based on expectations and on phrases built by bottom-up evidence exist in parallel and compete for activation.

Again, both linguistic and computational concerns support this approach. According to GB, phrase structure is projected from a lexical item; no $\overline{X}$ phrase can exist without being licensed in this way. Interpreting this principle computationally as a constraint on structure building, rather than one which checks already computed structure, increases the efficiency of the approach. Disallowing precomputation in the model not only limits the number of nodes that are created, but also simplifies the structure building, constraint satisfaction, and spreading activation algorithms significantly.

Not only do each of these decisions receive independent linguistic and computational support, these issues are in fact a set of interrelated choices. The motivation for integrating symbolic and numeric computation was presented above. By granting symbolic information a primary role in the processing of the network, the opportunity for building structure arises. Dynamically creating structure in turn discourages top-down precomputation. The fact that these are mutually constraining decisions yields a hybrid architecture with a coherent cluster of properties. The following section demonstrates that the combined result of these interdependent choices is a principled architecture from which the desired structural disambiguation behavior emerges.

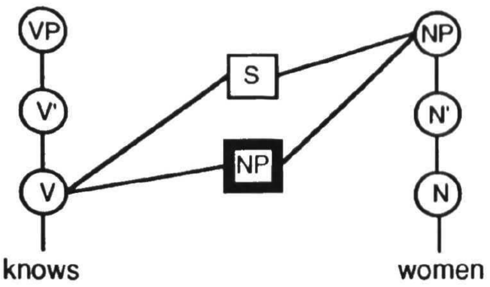
4 Predictions of the Model

Section 2 presented the two types of psycholinguistic observation which a theory of structural disambiguation should explain: the conflicting evidence concerning the degree to which structural alternatives are maintained in parallel, and the pattern of structural preferences which people exhibit. This section describes the behavior of the model relevant to these two issues. First, the critical properties which underlie the model's behavior will be presented. Next, the model's restricted parallelism will be seen to resolve the seemingly contradictory evidence for serial and parallel processing of structural alternatives. Finally, the preference behavior of the model on some illustrative cases of structural ambiguity will be discussed.

The property of the model which yields a unified theory of structural disambiguation is the process of active communication of symbolic and numeric information within the parsing network. When a phrase is created, the parser establishes nodes for the potential attachments of the phrase into the parse tree. The phrase must then actively communicate features to its neighbors in order to determine which attachments are valid. These attachment nodes are the structural alternatives for the phrase; their activation level encodes their relative preference. They receive numeric activation from their neighbors across weighted connections; these weights encode the frequency with which one phrase expects to attach to another. Activation decays over time if it is not reinforced. A competitive activation method (Reggia, Marsland, & Berndt, 1988) provides a focusing mechanism to sharpen the preference for an alternative.

Integrating serial and parallel behavior

Parallelism is restricted in the model by disallowing top-down precomputation. Thus a phrase which is actively seeking an attachment can only communicate with other structures which have received evidence from the input. In addition, a competitive mechanism required to control spreading activation attempts to focus activation on a single alternative. These two properties lead to observed serial behavior in the parser, accounting for the results of psycholinguistic experiments which support a serial mechanism. For example, the

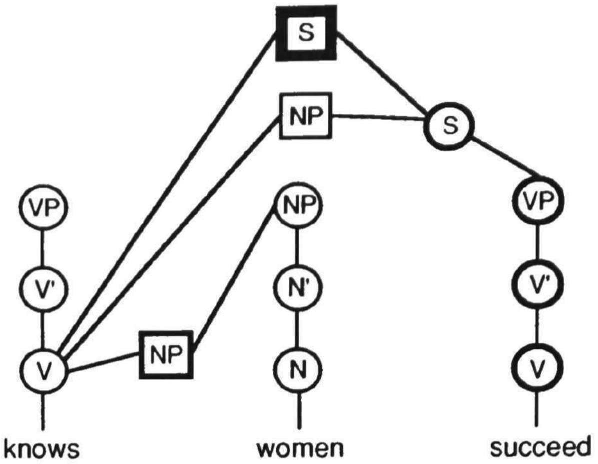


The NP may attach as the S object of *know* or as the NP object of *know*. Since the features of the first attachment expectation are incompatible with the NP, only the second attachment node remains active.

Figure 6: Initial attachment possibilities for *women*.

consistent preference for the NP–to–VP attachment in Figure 1 is simply due to the fact that it is the *only* attachment initially available for the NP. The parser has not encountered any overt evidence of a sentential object, so the NP can only make the attachment directly to the VP, as shown in Figure 6. In cases where multiple attachments are available, as was the case in Figure 2, the competitive activation mechanism will focus on one of them, also leading to seemingly serial behavior.[1] In either case, evidence that the initially preferred attachment is incorrect leads to a delay in processing, since the new structural alternative must compete for activation with the established attachment. This serial behavior mimics that of the human parser demonstrated in the analysis of eye-movements recorded while people read these types of temporarily ambiguous sentences (Frazier & Rayner, 1982).

This behavior is not inconsistent with a fundamentally parallel architecture, however. The model maintains multiple structural alternatives for which evidence exists, and projects active expectations. These expectations are particularly relevant in accounting for data showing that multiple alternatives are in some form available to the parser even when evidence from the input has not been encountered. Gorrell (1987), using a lexical decision task immediately following the NP, showed that the non-preferred structural alternative in Figure 1 (that is, the NP–to–S attachment) could prime a verb. Gorrell took this as clear evidence of precomputation of all possible structural alternatives. However, the active expectations of the model here can easily account for the observed behavior. Figure 7 shows that since the verb *know* actively expects either an NP or S argument, the S expectation primes the subsequent occurrence of a verb in the input.[2]



The verb's active attachment nodes encode the expectation of an NP or S object. The S expectation primes the S projected from the tense features of the verb.

Figure 7: Priming of a subsequent S.

Thus, even with the restriction of only maintaining the knowledge of alternative possibilities, rather than precomputing their structure, the model predicts the results of experiments supporting parallelism in situations of structural ambiguity. On the other hand, the restrictedness of the parallelism accounts for the serial effects displayed in experiments testing structural preferences. This explanation of the range of behaviors in processing structural ambiguities is the result of the inherent integration of aspects of serialism and parallelism in the model.

Unifying structural preferences

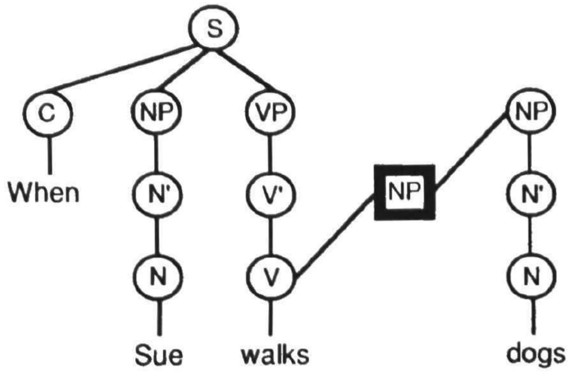
Fundamental properties of the model directly relate to the cognitive principles responsible for human structural preferences that were noted in Section 2. The need for a phrase to actively determine its valid attachments predicts that people will show a preference for attachments which allow immediate structuring of input. The decay of activation explains recency effects in structural preference, which correspond to a decrease in salience of older attachment sites. Weights increase salience by strengthening activation of more frequent alternatives. And finally, activation of expectations primes certain alternatives. These structural preference mechanisms interact in the resolution of each instance of structural ambiguity. A predictive model of structural disambiguation thus arises from these natural properties.

For example, the active attachment behavior of the model, in conjunction with the restricted parallelism discussed above, provides a simple account of many of the Minimal Attachment and Late Closure cases. Fig-

---

[1] The factors involved in focusing preferences will be discussed below.

[2] The verb in English carries tense features which trigger the building of an S node. More precisely, it is the attachment of this S node to the VP which is primed by the S expectation.



Only one attachment is initially available for the NP: as the NP object of *read*. (No main clause S node has been projected.)

Figure 8: Initial attachment possibilities for *dogs*.

ures 6 and 8 demonstrate that the alternative attachments are just not available at the time the NP is processed. Thus Minimal Attachment, represented in Figure 6, results not from an explicit comparison of the complexity of various choices as in previous models (for example, Frazier, 1987; Gorrell, 1987; Clark, 1988), but from the active nature of the model's syntactic phrases. Furthermore, cases of Late Closure as in Figures 8 are accounted for by the same properties of the model. That is, when the NP begins to actively seek an attachment, the parser has not yet projected a main clause S node, so the embedded VP is the only possible attachment site. This gives a uniform explanation of a range of structural preferences for which two attachment strategies were previously thought to be required.

The other standard cases of Late Closure, which were exemplified in Figure 2, are predicted by the recency effects which result from the decay of activation. Although both attachments are available, the higher attachment has less activation because that verb is more distant in the input and its activation has decayed. As the two verbs actively compete for the adverbial attachment, the more recent verb has the advantage of more activation and will "win" the competition. A similar effect arises from lexical frequencies, which are encoded by the weights on connections to possible arguments of a verb, and provide more or less advantage to potential attachments. (See Figure 5.) Priming in the form of an expectation also leads to the active advantage of an attachment, as discussed above and demonstrated in Figure 7.

Thus we have a model which accounts for a wide range of structural preferences with the single principled mechanism of active symbolic and numeric information. Not only does the model predict the various preferences, but it does so with a spreading activation mechanism which naturally integrates their interaction as well.

5 Related Approaches

Recent work of McRoy & Hirst (1989) similarly attempts to unify a broad range of syntactic influences on structural preferences. However, the timing effects they seek are not a natural result of their parsing architecture; in fact, they must explicitly build in the interaction of preferences and timing. Furthermore, their serial, race-based parser is unable to account for the parallel aspects of the processing of ambiguities, as is the similar model of Frazier & Rayner (1982). Cottrell's (1985) connectionist parser results in Minimal Attachment behavior due to the nature of spreading activation, but falls short of accounting for the related principle of Late Closure with the same mechanism. In the parallel models of Gorrell (1987) and Clark (1988), the ranking mechanisms proposed for determining structural preferences are not a fundamental aspect of the parsing architecture, and each fails to capture recency and lexical preference effects.

6 Conclusions

This paper has presented a hybrid massively parallel parsing architecture which integrates symbolic and numeric processing in a linguistically and computationally motivated way. Behavior mimicking human processing of structural ambiguities emerges from the inherent properties of this architecture. The serial aspects of spreading activation and the restriction on top-down precomputation provide a natural explanation for the seemingly irreconcilable range of serial and parallel behaviors in processing structural alternatives. The property of active symbolic and numeric information leads to a principled account of structural preferences, unifying with a single mechanism the effects of various previously proposed preference strategies, such as Minimal Attachment and Late Closure.

Acknowledgments

This work has been supported by University of Maryland Graduate School Fellowships, NSF Grant IST-8451430, and the University of Maryland Institute for Advanced Computing Studies (UMIACS).

Thanks to Amy Weinberg, James Reggia, and Sven Dickinson for helpful comments on earlier drafts of this paper.

References

Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa Lectures.* Dordrecht: Foris Publications.

Chomsky, N. (1986). *Barriers.* Cambridge: MIT Press.

Clark, R. (1988). "Parallel Processing and Local Optimization." Talk given at the University of Maryland Processing Workshop, December 9, 1988.

Cottrell, G. W. (1985). "Connectionist Parsing." *Proceedings of the Seventh Annual Conference of the Cognitive Science Society*, 201–211.

Ford, M., J. Bresnan, and R. Kaplan (1982). "A competence-based theory of syntactic closure." In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations.* Cambridge: MIT Press.

Frazier, L. (1978). *On Comprehending Sentences: Syntactic Parsing Strategies.* Unpublished doctoral dissertation, University of Connecticut. Distributed by Indiana University Linguistics Club.

Frazier, L. (1987). "Sentence processing." In M. Coltheart (Ed.), *Attention and Performance XII.* Hillsdale, NJ: LEA.

Frazier, L., and J. D. Fodor (1978). "The Sausage Machine: A new two-stage parsing model." *Cognition* **6**, 291–325.

Frazier, L., and K. Rayner (1982). "Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences." *Cognitive Psychology* **14**, 178–210.

Gorrell, P. (1987). "Structural ambiguity and syntactic priming: Toward a theory of ranked-parallel parsing." Manuscript, University of Maryland.

Hendler, J. (1986). Integrating Marker-Passing and Problem Solving: A Spreading Activation Approach to Improved Choice in Planning." Technical Report 1624, Computer Science Department, University of Maryland, College Park, Maryland.

Hendler, J. (1987). "Marker-passing and Microfeatures." *Proceedings of the Tenth International Joint Conferences on Artificial Intelligence*, 151–154.

Kimball, J. (1973). "Seven principles of surface structure parsing in natural language." *Cognition* **2**:**1**, 15–47.

McRoy, S. and G. Hirst (1989). "Race-Based Parsing and Syntactic Disambiguation." Manuscript, University of Toronto. Submitted for publication.

Nicol, J. (1988). "Coreference Processing During Sentence Comprehension: A Review of On-Line Research." Manuscript, University of Arizona.

Reggia, J., P. Marsland, and R. Berndt (1988). "Competitive Dynamics in a Dual-Route Connectionist Model of Print-to-Sound Transformation." *Complex Systems.*

Selman, G., and G. Hirst (1985). "A Rule-Based Connectionist Parsing Scheme." *Proceedings of the Seventh Annual Conference of the Cognitive Science Society*, 212–219.

Waltz, D., and J. Pollack (1985). "Massively parallel parsing: A strongly interactive model of natural language interpretation." *Cognitive Science* **9**, 51–74.