A View of Parsing

Ronald M. Kaplan
Xerox Palo Alto Research Center

The questions before this panel presuppose a distinction between parsing and interpretation. There are two other simple and obvious distinctions that I think are necessary for a reasonable discussion of the issues. First, we must clearly distinguish between the *static specification* of a process and its *dynamic execution*. Second, we must clearly distinguish two purposes that a natural language processing system might serve: one legitimate goal of a system is to perform some practical task efficiently and well, while a second goal is to assist in developing a scientific understanding of the cognitive operations that underlie human language processing. I will refer to parsers primarily oriented towards the former goal as Practical Parsers (PP) and refer to the others as Performance Model Parsers (PMP). With these distinctions in mind, let me now turn to the questions at hand.

1. The Computational Perspective.

From a computational point of view, there are obvious reasons for distinguishing parsing from interpretation. Parsing is the process whereby linearly ordered sequences of character strings annotated with information found in a stored lexicon are transduced into labelled hierarchical structures. Interpretation maps such structures either into structures with different formal properties, such as logical formulas, or into sequences of actions to be performed on a logical model or database. On the face of it, unless we ignore the obvious formal differences between string--to--structure and structure--to--structure mappings, parsing is thus formally and conceptually distinct from interpretation. The specifications of the two processes necessarily mention different kinds of operations that are sensitive to different features of the input and express quite different generalizations about the correspondences between form and meaning.

As far as I can see, these are simply factual assertions about which there can be little or no debate. Beyond this level, however, there are a number of controversial issues. Even though parsing and interpretation operations are recognizably distinct, they can be combined in a variety of ways to construct a natural language understanding system. For example, the static specification of a system could freely intermix parsing and interpretation operations, so that there is no part of the program text that is clearly identifiable as the parser or interpreter, and perhaps no part that can even be thought of as more parser-like or interpreter-like than any other. Although the microscopic operations fall into two classes, there is no notion in such a system of separate parsing and interpretation components at a macroscopic level. Macroscopically, it might be argued, a system specified in this way does not embody a parsing/interpretation distinction.

On the other hand, we can imagine a system whose static specification is carefully divided into two parts, one that only specifies parsing operations and expresses parsing generalizations and one that involves only interpretation specifications. And there are clearly untold numbers of system configurations that fall somewhere between these extremes.

I take it to be uncontroversial that, other things being equal, a homogenized system is less preferable on both practical and scientific grounds to one that naturally decomposes. Practically, such a system is easier to build and maintain, since the parts can be designed, developed, and understood to a certain extent in isolation, perhaps even by people working independently. Scientifically, a decomposable system is much more likely to provide insight into the process of natural language comprehension, whether by machines or people. The reasons for this can be found in Simon's classic essay on the *Architecture of Complexity*, and in other places as well.

The debate arises from the contention that there are important "other things" that cannot be made equal, given a completely decomposed static specification. In particular, it is suggested that parsing and interpretation operations must be partially or totally interleaved during the execution of a comprehension process. For practical systems, arguments are advanced that a "habitable" system, one that human clients feel comfortable using, must be able to interpret inputs before enough information is available for a complete syntactic structure or when the syntactic information that is available does not lead to a consistent parse. It is also argued that interpretation must be performed in the middle of parsing in the interests of reasonable efficiency: the interpreter can reject sub-constituents that are semantically or pragmatically unacceptable and thereby permit early truncation of long paths of syntactic computation. From the performance model perspective, it is suggested that humans seem able to make syntactic, semantic, and pragmatic decisions in parallel, and the ability to simulate this capability is thus a condition of adequacy for any psycholinguistic model.

All these arguments favor a system where the operations of parsing and interpretation are interleaved during dynamic execution, and perhaps even executed on parallel hardware (or wetware, from the PMP perspective). If parsing and interpretation are run-time indistinguishable, it is claimed, then parsing and interpretation must be part and parcel of the same monolithic process.

Of course, whether or not there is dynamic fusion of parsing and interpretation is an empirical question which might be answered differently for practical systems than for performance models, and might even be answered differently for different practical implementations. Depending on the relative computational efficiency of parsing versus interpretation operations, dynamic interleaving might increase or decrease overall system effectiveness. For example, in our work on the LUNAR system (Woods, Kaplan, & Nash-Webber, 1972), we found it more efficient to defer semantic processing until after a complete, well-formed parse had been discovered. The consistency checks embedded in the grammar could rule out syntactically unacceptable structures much more quickly than our particular interpretation component was able to do. More recently, Martin, Church, and Ramesh (1981) have claimed that overall efficiency is greatest if all syntactic analyses are computed in breadth-first fashion before any semantic operations are executed. These results might be taken to indicate that the particular semantic components were poorly conceived and implemented, with little bearing on systems where interpretation is done "properly" (or parsing is done improperly). But they do make the point that a practical decision on the dynamic fusion of parsing and interpretation cannot be made a priori, without a detailed study of the many other factors that can influence a system's computational resource demands.

Whatever conclusion we arrive at from practical considerations, there is no reason to believe that it will carry over to performance modelling. The human language faculty is an evolutionary compromise between the requirements that language be easy to learn, easy to produce, and easy to comprehend. Because of this, our cognitive mechanisms for comprehension may exhibit acceptable but not optimal efficiency, and we would therefore expect a successful PMP to operate with psychologically appropriate *in*efficiencies. Thus, for performance modelling, the question can be answered only by finding cases where the various hypotheses make crucially distinct predictions concerning human capabilities, errors, or profiles of cognitive load, and then testing these predictions in a careful series of psycholinguistic experiments. It is often debated, usually by non-linguists, whether the meta-linguistic intuitions that form the empirical foundation for much of current linguistic theory are reliable indicators of the native speaker's underlying competence. When it comes to questions about internal processing as opposed to structural relations, the psychological literature has demonstrated many times that intuitions are deserving of even much less trust. Thus, though we may have strong beliefs to the effect that parsing and interpretation are psychologically inseparable, our theoretical commitments should rather be based on a solid experimental footing. At this point in time, the experimental evidence is mixed: semantic and syntactic processes are interleaved on-line in many situations, but there is also evidence that these processes have a separate, relatively non-interacting run-time course.

However, no matter how the question of dynamic fusion is ultimately resolved, it should be clear that dynamic interleaving or parallelism carries no implication of static homogeneity. A system whose run-time behavior has no distinguishable components may nevertheless have a totally decomposed static description. Given this possibilty, and given the evident scientific advantages that a decomposed static specification affords, I have adopted in my own research on these matters the strong working hypothesis that a statically decomposable system *can* be constructed to provide the necessary efficiencies for practical purposes and yet, perhaps with minor modifications and further stipulations, still support significant explanations of psycholinguistic phenomena.

In short, I maintain the position that the "true" comprehension system will also meet our pre-theoretic notions of scientific elegance and "beauty". This hypothesis, that truth and beauty are highly correlated in this domain, is perhaps implausible, but it presents a challenge for theory and implementation that has held my interest and fascination for many years.

2. The Linguistic Perspective.

While it is certainly true that our tools (computers and formal grammars) have *shaped* our views of what human languages and human language processing may be like, it seems a little bit strange to think that our views have been *warped* by those tools. Warping suggests that there is some other, more accurate view that we would have come to either without mathematical or computational tools or with a set of formal tools with a substantially different character. There is no way in principle to exclude such a possibility, but it could be that we have the tools we have because they harmonize with the capabilities of the human mind for scientific understanding. That is, athough substantially different tools might be better suited to the phenomena under investigation, the results derived with those tools might not be humanly appreciable. The views that have emerged from using our present tools might be far off the mark, but they might be the only views that we are capable of.

Perhaps a more interesting statement can be made if the question is interpreted as posing a conflict between the views that we as *computational* linguists have come to, guided by our present practical and formal understanding of what constitutes a reasonable computation, and the views that theoretical linguists, philosophers, and others similarly unconstrained by concrete computation, might hold. Historically, computational grammars have represented a mixture of intuitions about the significant structural generalizations of language and intuitions about what can be processed efficiently, given a particular implementation that the grammar writer had in the back of his or her mind.

This is certainly true of my own work on some of the early ATN grammars. Along with many others, I felt an often unconscious pressure to move forward along a given computational path as long as possible before throwing my grammatical fate to the parser's general nondeterministic choice mechanisms, even though this usually meant that register contents had to be manipulated in linguistically unjustified ways. For example, the standard ATN account of passive sentences used register operations to avoid backtracking that would reanalyze the NP that was initially parsed as an active subject. However, in so doing, the grammar confused the notions of surface and deep subjects, and lost the ability to express generalizations concerning, for example, passive tag questions.

In hindsight, I consider that my early views were "warped" by both the ATN formalism, with its powerful register operations, and my understanding of the particular top-down, left-right underlying parsing algorithm. As I developed the more sophisticated model of parsing embodied in my General Syntactic Processor, I realized that there was a systematic, non-grammatical way of holding on to functionally mis-assigned constituent structures. Freed from worrying about exponential constituent structure nondetermism, it became possible to restrict and simplify the ATN's register operations and, ultimately, to give them a non-procedural, algebraic interpretation. The result is a new grammatical formalism, Lexical-Functional Grammar (Kaplan & Bresnan, in press), a formalism that admits a wider class of efficient computational implementations than the ATN formalism just because the grammar itself makes fewer computational commitments. Moreover, it is a formalism that provides for the natural statement of many language particular and universal generalizations. It also seems to be a formalism that facilitates cooperation between linguists and computational linguists, despite their differing theoretical and methodological biases.

Just as we have been warped by our computational mechanisms, linguists have been warped by their formal tools, particularly the transformational formalism. The convergence represented by Lexical-Functional Grammar is heartening in that it suggests that imperfect tools and understanding can and will evolve into better tools and deeper insights.

3. The Interactions.

As indicated above, I think computational grammars have been influenced by the algorithms that we expect to apply them with. While difficult to weed out, that influence is not a theoretical or practical necessity. By reducing and eliminating the computational commitments of our grammatical formalism, as we have done with Lexical-Functional Grammar, it is possible to devise a variety of different parsing schemes. By comparing and contrasting their behavior with different grammars and sentences, we can begin to develop a deeper understanding of the way computational resources depend on properties of grammars, strings, and algorithms. This understanding is essential both to practical implementations and also to psycholinguistic modelling. Furthermore, if a formalism allows grammars to be written as an abstract characterization of string--structure correspondences, the grammar should be indifferent as to recognition or generation. We should be able to implement feasible generators as well as parsers, and again, shed light on the interdependencies of grammars and grammatical processing.

Let me conclude with a few comments about the psychological validity of grammars and parsing algorithms. To the extent that a grammar correctly models a native speaker's linguistic competence, or, less tendentiously, the set of meta-linguistic judgments he is able to make, then that grammar has a certain psychological "validity". It becomes much more interesting, however, if it can also be embedded in a psychologically accurate model of speaking and comprehending. Not all competence grammars will meet this additional requirement, but I have the optimistic belief that such a grammar will someday be found.

It is also possible to find psychological validation for a parsing algorithm in the absence of a particular grammar. One could in principle adduce evidence to the effect that the architecture of the parser, the structuring of its memory and operations, corresponds point by point to well-established cognitive mechanisms. As a research strategy for arriving at a psychologically valid model of comprehension, it is much more reasonable to develop linguistically justified grammars and computationally motivated parsing algorithms in a collaborative effort. A model with such independently motivated yet mutually compatible knowledge and process components is much more likely to result in an explanatory account of the mechanisms underlying human linguistic abilities.

References

Kaplan, R. & Bresnan, J. Lexical-functional grammar: A formal system for grammatical representation. In J. Bresnan (ed.), *The mental representation of grammatical relations*. Cambridge: MIT Press, in press.

Martin, W., Church, K., & Ramesh, P. Paper presented to the Symposium on Modelling Human Parsing Strategies, University of Texas at Austin, March, 1981.

Woods, W., Kaplan, R., & Nash-Webber, B. *The Lunar sciences natural language information system*. Cambridge: Bolt Beranek and Newman, Report 2378, 1972.