Principle-Based Parsing: A Symposium

Robert C. Berwick
MIT Artificial Intelligence Laboratory

Principle-based parsing, sometimes called principles-and-parameters parsing, is a relatively new approach to language analysis that replaces the traditionally large set of rules used to analyze sentences on a language-by-language basis with a much smaller, fixed set of parameterized, universal principles. The principles interact deductively to replace many rules. In such a system there is no "rule" of passive, encoded perhaps in context-free rule, if-then format, or a transition network; instead, there are broader axioms from which one can "deduce" the possibility of a passive-type sentence. Thus the shift is much like that has occurred in expert systems from a shallow if-then form to an explanatory system that reasons from first principles. The underlying linguistic theory is most often associated with the so-called "principles and parameters" framework or government-binding (GB) theory of linguistic analysis.

But this shift from rules to principles raises many serious questions for cognitive science. Is principle-based parsing possible? Can it really work universally, for many different languages, thus raising the possibility of machine translation once again? How can efficient parsing be done with general principles? What implementations are possible? How do alternative machine architectures fit in, including connectionism? In the five symposium papers that follow, and a general discussion we hope to address these and other problems in this rapidly developing area of natural language processing.

Principle-Based Parsing

Steven P. Abney
Bell Communications Research

I consider characterizations of principle-based parsing or government-binding theory parsing (GB parsing) that have been made in the literature. A common theme is the *transparency hypothesis*, according to which there is a close correspondence between the theoretical constructs of GB syntactic theory and either the data objects manipulated by the parser, or the procedures that make up the parser. There are at least three different motivations for the transparency hypothesis: (1) because it gives substance to the conviction that syntactic constructs represent real objects, and are not just convenient fictions, (2) because it makes program verification easier, (3) because it serves as a source of inspiration for new parsing techniques.

Motivation (1) is unjustified. There are many objects in the domain of parsing that are real, in the sense of providing the best explanation of surface phenomena, but which are neither data objects nor procedures. For example, in an LR parser generated from a grammar G, there is no data object or procedure that can be identified as a rule of G; the rules of G exist only in their combined consequences for the LR tables. Nonetheless, the best explanation of the behavior of the parser is in terms of the rules of G.

Motivation (2) is valid, but program verification is only one concern of many in software engineering. If (2) is our only motivation, then current work in principle-based parsing accords disproportionate weight to program verification.

Motivation (3) can hardly be faulted, though it makes the transparency hypothesis of value only for those people whom it inspires. I consider some GB-inspired techniques that have value independent of their relation to GB.

Principle-Based Machine Translation

Bonnie J. Dorr
MIT Artificial Intelligence Laboratory

This talk will describe UNITRAN, an implemented principle-based machine translation system. The task of cross-linguistic translation is difficult because there are several types of phenomena within any given language; moreover, the number of ways in which these phenomena can be exhibited is potentially enormous across different source and target languages. The primary characteristic of UNITRAN is that it operates uniformly across English, Spanish, and German, while still accounting for knowledge that is specific to each language.
Consider the following translation from English to Spanish:

John broke into the room ⇒ Juan forzó la entrada al cuarto

In this example, the source-language sentence diverges both syntactically and lexically from the target-language sentence. The syntactic divergence shows up in the translation of the single-object construction *into the room* into the double-object construction *la entrada al cuarto*. The lexical divergence shows up in the realization of the simple verb *break* as the composite form *forzar la entrada* (literally, *force entry*). The UNITRAN system solves these types of divergences by making use of two levels of processing: a syntactic level based on principles and parameters, and a lexical-semantic level based on the compositionality of lexical items.

UNITRAN differs from other translation systems in that it is able to maintain language-independent information while still processing many types of language-specific phenomena associated with different source and target languages. Within this framework, language-independent conceptual representations are used as the pivot between the source and target languages, and language-specific phenomena are accounted for by means of parameter settings and lexically-specified information.

The Computational Implementation of Principle-based Parsers

Sandiway Fong
MIT Artificial Intelligence Laboratory

Recently, there have been some interest in the implementation of grammatical theories based on the principles and parameters approach. In this framework, a fixed set of universal principles interact deductively to account for diverse linguistic phenomena.

In this talk, I will address the issue of how to organize such principles for efficient processing. That is, what is an appropriate 'machine architecture' for parsers based on the principles-and-parameters framework? I will discuss two possible approaches. First, I will discuss the design of parsers that build complete phrase structure representations and then apply principles in order to rule out those that are ill-formed. In this approach, I will discuss whether it is possible to exploit the relatively-free interaction of linguistic principles to engineer more efficient parsers. In particular, I will show that the hitherto largely ignored issue of principle ordering can have a large effect on parsing efficiency, and hence, should be of much concern to parser designers.

An alternative approach is to interleave the task of building phrase structure representations with principle application. I will address the question: Is principle interleaving a more effective parsing strategy than a non-interleaved approach? I will discuss both the conditions under which interleaving may be a more effective strategy, and circumstances where a non-interleaved approach may be preferable.

I will present the results of various experiments on different machine configurations in the context of an implemented system that was built to investigate these issues. This system allows the automatic construction of a whole family of parsers with different architectures that all use the same underlying set of principles.

Constraint-oriented Principle-based Parsing

Mark Johnson
Brown University

One of the difficulties in the construction of computational models for Chomsky's "Government and Binding" theory is that the appropriate representations of an utterance at different linguistic levels are determined by a system of tightly interacting principles. Simple-minded implementations which construct these levels of representation one by one cannot fully exploit the constraints on the first levels of representation they construct imposed by other as yet unconstructed levels: a disastrous short-coming since these reduce an infinite search space to a finite one. This problem can be avoided by determining in advance some or all of the constraints that the later, as yet unconstructed levels of representations will impose the other levels. It leads to a constraint-oriented rather than representation-oriented perspective: the construction of one representation requires only the constraints originating from the other representations, not the representations themselves. If *all* constraints originating from some level of representation can be determined without that level of representation, then that level need not be constructed by the computational model. Simple natural language parsers will be described that function in exactly this way, avoiding the explicit construction of several levels of representation, although provably producing analyses that satisfy all constraints on all linguistic levels (including those unconstructed).

Such computational models pose interesting problems for the "mental representation as data structure" analogy common in cognitive science, since arguably the computational model does exploit all of the levels of representations of an utterance (where else could the relevant constraints come from?), even though it does not construct data structures corresponding to each level of representation.

Relaxation Techniques for Principle-Based Parsing

Edward P. Stabler, Jr.
UCLA Dept of Linguistics

Consider parsers which are "principle-based" in the following sense: structural representations are obtained by deductive reasoning from an explicitly represented linguistic theory. Such parsers can still differ along many dimensions. For linguists and psychologists interested in the bearing of linguistic theory on accounts of human parsing, the differences between the following principle-based views are significant, where Con(G) is the set of consequences of grammar G:

> (G) Parsing from first principles: the linguists' grammar G is represented and used directly
>
> (EG) Parsing from equivalent principles: using a theory EG such that Con(EG)=Con(G)
>
> (WG) Parsing from weaker principles: using a theory WG such that Con(WG) is a proper subset of Con(G)
>
> (DG) Parsing from different principles: using a theory DG such that Con(D) is not a subset of Con(G).

Focusing on government-and-binding theory, it will be argued that (G) has serious complexity problems that can be avoided by a logically equivalent, "relaxed" approach (EG), which solves parsing problems after first solving appropriate approximations. A version of (EG) together with a plausible performance story would remove the motivation for going to (WG); and (DG) is the least appealing option, only of interest if the other options fail, and they haven't. I will conclude by considering whether the presented argument against (G) depends on assumptions that have been effectively challenged by connectionist approaches to constraint satisfaction.