Squibs and Discussions

Ambiguity-preserving Generation with LFG- and PATR-style Grammars

Jurgen Wedekind*
University of Stuttgart

Ronald M. Kaplan†
Xerox Palo Alto Research Center

The widespread ambiguity of natural language presents a particular challenge for machine translation. The translation of an ambiguous source sentence may depend on first determining which reading of the sentence is contextually appropriate and then producing a target sentence that accurately expresses that reading. This may be difficult or even impossible to accomplish when resolution of the source ambiguity depends on a complete understanding of the text, or when several readings are contextually appropriate. An attractive alternative strategy is to circumvent the need for disambiguation by generating a target sentence that has exactly the same ambiguities as the source. In this brief note we investigate whether ambiguity-preserving generation is possible when syntactic structures are described by the mechanisms of LFG- or PATR-style grammars (Kaplan and Bresnan 1982, Shieber et al. 1983). Mechanisms of this sort associate attribute-value structures with trees derived in accordance with a context-free grammar. Our result also applies to other systems such as HPSG (Pollard and Sag 1994) whose formal devices are powerful enough to simulate, albeit indirectly, the effect of context-free derivation.

Consider as an example the well-known ambiguous sentence (1)

(1) John saw the man with the telescope.

for which some LFG or PATR grammar might provide alternative f-structures equivalent to the more compact predicate-calculus formulas indicated in (2).

(2) a. *with_the_telescope(see(John, man))*

b. *see(John, with_the_telescope(man))*

The problem of translating this sentence appropriately into, say, German could be handled by disambiguating its parsing result (2) (i.e., choosing one of these structures/formulas), converting that to an appropriate German f-structure (or leaving it alone if the result can serve as an interlingua), and then generating a German sentence that would have that meaning as (hopefully, the only) one of its interpretations. Disambiguation is the major formal obstacle in this approach, since parsing algorithms exist when the grammatical formalisms are off-line parsable (Kaplan and Bresnan 1982), and the generation problem is known to be decidable even without the off-line parsable restriction (Wedekind 1995). However, for this sentence it should be possible to side-step the disambiguation problem because there is a German sentence (3) that expresses exactly the same ambiguity as the original English.

* Institute for Natural Language Processing, University of Stuttgart, Azenbergstr. 12, D-70174 Stuttgart, Germany. E-mail: juergen@ims.uni-stuttgart.de

† Xerox PARC, 3333 Coyote Hill Road, Palo Alto, CA 94304, USA. E-mail: kaplan.parc@xerox.com

(3) Hans sah den Mann mit dem Fernrohr.

Unfortunately, an ambiguity-preserving translation does not exist for all source sentences. This is illustrated by the English sentence (4), with German taken again as the target language.

(4) The duck is ready to eat.

The two readings of (4) are given in (5).

(5) a. *ready(duck, eat(someone, duck))*

b. *ready(duck, eat(duck, something))*

These interpretations have to be expressed in separate German sentences, as in (6).

(6) a. Die Ente kann jetzt gegessen werden.

b. Die Ente ist zum Fressen bereit.

Preservable and unpreservable ambiguity can occur in complex patterns. The sentence (7)

(7) John saw her duck with the telescope.

has at least the four readings indicated in (8).

(8) a. *with_the_telescope(see(John, her_duck))*

b. *see(John, with_the_telescope(her_duck))*

c. *with_the_telescope(see(John, duck(her)))*

d. *see(John, with_the_telescope(duck(her)))*

No single sentence in German expresses all these readings, but accurate translation for this case does not require a full disambiguation. The readings (8a,b) can be expressed by (9a) and the readings (8c,d) by (9b).

(9) a. Hans sah ihre Ente mit dem Fernrohr.

b. Hans sah sie mit dem Fernrohr untergehen.

Thus only "duck" has to be disambiguated—the PP-attachment ambiguity is preserved in both translations.

We see from these examples that the costly and difficult process of disambiguation can be avoided in some circumstances but is necessary for accurate translation when an ambiguity-preserving target construction does not exist. The performance of a system may be improved, then, if the disambiguation process is initiated only when it has been determined that no target sentence can be generated that expresses exactly the set of readings found in the source. In this note we consider whether or not it is possible to make this kind of determination, and we arrive at an essentially negative result: the problem of ambiguity-preserving generation (and thus ambiguity-preserving translation) is unsolvable even if the languages are described by unification grammars for which the parsing and generation problems separately are computable.

Since the proof of this assertion is so simple, we can dispense almost entirely with preliminary formalizations. We need only the fact that an LFG- or PATR-style unification grammar G defines a binary relation Δ_G between terminal strings w and f-structures Φ as given in (10)

(10) $\Delta_G(w, \Phi)$ iff w is derivable with Φ according to G.

On the basis of Δ we can then show that ambiguity-preserving generation is undecidable.

Theorem
Let G be an arbitrary unification grammar and let $\{\Phi_1, \ldots, \Phi_l\}$ $(l > 1)$ be an arbitrary set of feature structures. Then it is undecidable whether there is a terminal string w such that $\Delta_G(w, \Phi_1) \wedge \cdots \wedge \Delta_G(w, \Phi_l)$.

Proof
We prove the theorem by reducing the problem to the emptiness problem of the intersection of arbitrary context-free languages, a problem that is known to be undecidable. Let G^1 and G^2 be two arbitrary context-free grammars whose nonterminal vocabularies, terminal vocabularies, start-symbols, and rules are given by $\langle V_N^1, V_T^1, S^1, R^1 \rangle$ and $\langle V_N^2, V_T^2, S^2, R^2 \rangle$, respectively. Without loss of generality we suppose further that $V_N^1 \cap V_N^2 = \emptyset$. On the basis of G^1 and G^2 we construct a unification grammar $G = \langle V_N, V_T, S, R \rangle$ with

$$\begin{aligned} V_N &= V_N^1 \cup V_N^2 \cup \{S\} \text{ and } S \notin V_N^1 \cup V_N^2 \\ V_T &= V_T^1 \cup V_T^2 \\ R &= R^1 \cup R^2 \cup \left\{ S \rightarrow \begin{array}{c} S^1 \\ (\uparrow \text{A}) = 1 \end{array}, \quad S \rightarrow \begin{array}{c} S^2 \\ (\uparrow \text{A}) = 2 \end{array} \right\}. \end{aligned}$$

By this construction the problem of whether there is a terminal string w with $\Delta_G(w, [\text{A } 1])$ and $\Delta_G(w, [\text{A } 2])$ reduces to the undecidable problem whether $L(G^1) \cap L(G^2) = \emptyset$. This is because all strings in $L(G^1)$ are assigned the f-structure [A 1], all strings in $L(G^2)$ are assigned [A 2], and only strings in the intersection are derived ambiguously with [A 1] and [A 2]. □

As a consequence of this theorem we know that we cannot appeal to a general algorithm for solving the problem of ambiguity-preserving translation. This does not rule out the possibility that solutions can be found for specific constructions in translating between certain language pairs—for example, the PP ambiguity between English and German—but these solutions may depend on a detailed, non-algorithmic contrastive analysis for those constructions and languages.

It is also possible that natural language grammars belong to a restricted subclass of the LFG and PATR formalisms with properties that do not support the particular proof we have given. The f-structures assigned by our grammar G are structurally unrelated to the strings they are assigned to, and this seems quite unrealistic. As a minimum, it seems that there should be some relationship, perhaps a simple proportion, between the size of an f-structure and the length of any string it is assigned to, and that such a relationship would reduce the problem to an intersection of finite sets. Further study is needed to determine which, if any, intuitively plausible restrictions will permit the computation of ambiguity-preserving generation in a way that is both effective and efficient.

References

Kaplan, R. and J. Bresnan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In J. Bresnan, editor, *The Mental Representation of Grammatical Relations*. The MIT Press, Cambridge, MA, pages 173–281.

Pollard, C. and I. Sag. 1994. *Head-Driven Phrase Structure Grammar*. The University of Chicago Press, Chicago.

Shieber, S., H. Uszkoreit, F. Pereira, J. Robinson, and M. Tyson. 1983. The formalism and implementation of PATR-II. In B. Grosz and M. Stickel, editor, *Research on Interactive Acquisition and Use of Knowledge*. SRI Final Report 1894. SRI International, Menlo Park, pages 39–79.

Wedekind, J. 1995. Some remarks on the decidability of the generation problem in LFG- and PATR-style unification grammars. In *Proceedings of the 7th Conference of the European Chapter of the Association for Computational Linguistics*. Dublin, pages 45–52.