THE FULCRUM APPROACH TO MACHINE TRANSLATION

Christine A. Montgomery
Language Systems Inc.
5959 Topanga Canyon Blvd.
Woodland Hills, CA 91367
chris@lsi.com
(818) 703-5034

In a paper from a distinguished collection of papers prepared for a 1959 course entitled "Computer Programming and Artificial Intelligence," Paul Garvin described two types of machine translation problems "in terms of the two components of the term: machine problems, and translation problems." While the machine problems made us crazy, the translation problems made us think differently about language than we might otherwise have done, which has had some advantages and some disadvantages in the long run. I will save anecdotes about the former and comments about the latter for the discussion.

In this paper I will focus on the translation problems and, in particular, the translation approach that was developed by Paul Garvin, with whom I was associated, initially at Georgetown University, and later in the Synthetic Intelligence Department of the Ramo-Wooldridge Corporation and successor corporations: Thompson Ramo Wooldridge and Bunker-Ramo.

The initial MT experiments at Georgetown University in 1954 were designed by Garvin and carried out jointly with IBM, who implemented Garvin's design on an IBM 701 computer. These experiments were necessarily quite limited in scope and were a demonstration of possibility rather than feasibility.

The approach Garvin developed subsequently, entitled "The Fulcrum Approach to Machine Translation," was based on what he referred to as the "fulcra" of syntactic phrases of various types. The fulcrum of a syntactic unit is the constituent that "provides the best point of leverage for its translation," essentially because it governs the syntactic behavior of the unit, functioning as the head of the given syntactic unit or phrase. Thus the verb or other predicate is the fulcrum of the clause, the noun is the fulcrum of the noun phrase, the participle or other governing modifier controls the participial or modifier phrase, the preposition governs the prepositional phrase. Syntactic processing of a source language sentence from a text was performed via a series of analytical passes through the sentence, in which the fulcra of the various types of syntactic phrases were determined, and the constituents of the given phrases identified.

Although we considered the syntactic passes to be the most important, they were outnumbered by a series of pre- and post-syntactic analysis passes through the sentence, as exemplified by the following list of topics discussed under "Illustration of [Translation] Process" in the initial section of of a Ramo-Wooldridge technical report submitted to Rome Air Development Center in February, 1961:

> Input processing of Cyrillic text
> Transliteration
> Dictionary lookup
> Morphological analysis of missing words
> "Reinflection" of English translations of missing words
> Idiom translation
> Syntactic analysis

Translation of "word combinations"
Article insertion
Insertion of prepositions for translation of case suffixes
Insertion of auxiliaries
Rearrangement of word order within clauses
Rearrangement of word order within modifier phrases
Rearrangement of word order within verb phrases

Many of the above processes appear quite trivial, but they were necessary to deal with the variety of phenomena making up the elements of typical Russian scientific texts, in order to translate them into English. An example of elements that had to be isolated in input processing were bracketed items called "number/symbol strings." These could be equations, quantities, or other expressions that might have a key syntactic role in the sentence, but they might also contain internal periods or other punctuation that could potentially result in erroneous identifications of phrase, clause, or sentence boundaries with occasionally catastrophic results, such as causing the text buffers to overflow into other memory areas. In at least one case this resulted in overwriting the operating system code and derailing not only our MT processing but the operations of the service bureau where our processing was performed, making us extremely unpopular clients there.

In the initial Fulcrum system described in the above report, analysis of the source sentence and synthesis of the target sentence were interleaved. Russian words were replaced by English equivalents (or sets of equivalents) during dictionary lookup, and subsequent processes were aimed at refining the raw, word-for-word translation produced as a result of the initial lexical substitution. Proper English word order (or an approximation of this) was produced by "rearrangement routines" which operated within specific types of syntactic phrases to re-sequence constituents according to the requirements of English word order.

At the time the infamous ALPAC report appeared, Garvin and his staff at Bunker-Ramo had just completed the design and partial implementation of Fulcrum II, in which analysis of the Russian source language sentence and synthesis of the English target language translation were viewed as separate, sequential processes. The Fulcrum II dictionary entries were substantially enriched, including more detailed syntactic and semantic information. These changes improved the attachment of governed structures and facilitated the revision of attachments based on subsequent analytical processing. Unfortunately, full implementation of this more flexible system was never achieved because of the well-known impact of the ALPAC report on MT funding.

REFERENCES

Bunker-Ramo Corporation. 1965. A Syntactic Analyzer Study. Final Technical Report (Contract AF30(602)-3506: Rome Air Development Center). 31 July 1965.

Garvin, Paul L. 1959. Machine Translation. In *Computer Programming and Artificial Intelligence.* University of Michigan, pp. 1-8.

Ramo-Wooldridge. 1961. Machine Translation Studies of Semantic Techniques. Final Technical Report. (Contract AF30(602)-3506: Rome Air Development Center). 22 February 1961.