Logos: the intelligent translation system

Ralph E. Hawes

Vice-President, Marketing, Logos Corporation, Waltham, Massachusetts, USA

Logos is stated to be a full-capability (lexical, syntactic, semantic) MT system on a standard word processor; one of the first commercial examples of artificial intelligence; and a powerful servant for a translator who can type. All functions are controlled from the translator's workstation, including the infinitely expandable dictionary, to which the user, in dialogue with the system, adds the semantic data necessary for idiomatic translation.

I would like to introduce our company - Logos Corporation - and its product, Logos™, the Intelligent Translation System™. I know that Logos is not a complete stranger to those of you who have been attending these conferences during the past several years. I searched past proceedings and came across a few brief references to it, but usually there were few details. I think people saw us as a special projects company, which had written translation programs for English to Vietnamese and English to Farsi.

Our work in these and other languages like French, Russian and German helped us to develop our product and our company, which has become much more than a company for special language projects.

Logos recognised that our hard-earned expertise in natural language processing coupled with the past decades' giant technological advances in computer price/performance provided us with a unique opportunity: to bring together a number of proven, powerful translation techniques plus some special ideas of our own, in a hardware-and-software environment that is economical and friendly. This blend of ingredients has resulted in our offering Logos, the

Intelligent Translation System, on Wang, which we believe to be the world's leading word processing system. The topics that I will cover are:

- The company - Logos Corporation
- The characteristics of a desirable language translation system
- Logos, the Intelligent Translation System.

THE LOGOS CORPORATION

Ever since the ALPAC report in 1966, which discouraged efforts to produce workable machine translation (MT) of natural language, workers in MT have started only a few new projects (that is, until very recently). In general, their results have been widely reported. But one of the largest, and most sustained, efforts in this field has, until now, received little attention: the effort of Logos Corporation.

Logos Corporation is an American company with a German subsidiary. Corporate and marketing headquarters are in the US, along with a development staff of about seventy people. Our current sales efforts are conducted from our sales office in Frankfurt, West Germany.

The company was founded in 1969 by Bernard E. Scott, who is its president and principal linguist. Supported by private capital and development contracts, the company has always pursued a single goal: workable machine translation .

In the 1970s, the company's best-known technical project was the production of English-to-Vietnamese translation for the US government. Subsequent projects entailed work in French, Russian, Farsi, Spanish and German. When we directed our efforts to developing the commercial Logos product, we were able to apply the lessons learned in these prior efforts, and concentrate on incorporating proved techniques while at the same time avoiding many traps and pitfalls.

All our years of experience have contributed to the specification and development of the Intelligent Translation System, which incorporates a language-independent Automatic Translator™ of advanced capabilities.

We believe the Logos system is the world's most advanced translation system, embodying a state-of-the-art translation program free from architectural or logical limitations, with unprecedented capability for translating accurately.

A DESIRABLE LANGUAGE TRANSLATION SYSTEM

Let's briefly examine the most desirable characteristics of a robust automated language translation system. To begin with:

- The system should be capable of translating on its own, at computer speed, without operator intervention.

- The system should be able to accept text to be translated in an economic, reliable, friendly manner. There should be several input options ranging from keyboard to floppy disk or tape and including OCR and telecommunications. This also means the ability to interface with today's most common computer equipment vendors both in teleprocessing protocol and in text format.

- The system should faithfully reproduce the format of the source text in the target text.

- The system should be capable of considering syntax, and the deep semantic issues involved in transferring meaning and nuance from one language to another - much more than just word equivalents.

- The system should have multi-target capability, able to translate from a single source to a number of targets.

- The system should be capable of learning new vocabulary, new semantically desirable transfers, and the technical jargon of new technologies.

- The system should be capable of unlimited logical expansion in any direction. Dictionary transfers can be projected to millions of words, and semantic tables to many, many rules for each or any word.

- The system should be able to translate the full range of expository textual material, technical documentation, and all types of reports and factual articles. It should not require input with an artificially restricted vocabulary.

- There should be a technique wherein both users and the developer can improve the system through updates. An ideal system should not become a custom, stand-alone system soon after it is installed. While it must be able to incorporate the special requirements of

each user so as to make them appear an integral part of the system, these adaptations must not prevent system enrichment from a central source.

- The system must be cost-effective when compared with present manual efforts or with other MT systems.

- The   system   should be friendly and easy to use in all aspects.

LOGOS, THE INTELLIGENT TRANSLATION SYSTEM

In 1982, at Germany's Hanover Fair, the company announced its first commercial product: the Intelligent Translation System, with the Automatic Translator running in a microprocessor. The operating environment was and is the hardware and software of the Wang OIS 140/145 Office Information System. This has recently been extended to include the larger Wang VS System.

Our first language pair was German-to-English, which we are demonstrating during this meeting. At the 1983 Hanover Fair, the company announced English-to-German and demonstrated that product, which is currently in field test.

From a customer's point of view, this product is a language translator that operates on a word processor - and a standard word processor at that. And the market response has been enthusiastic, as you might expect.

Let's examine this product in light of the desirable characteristics previously described...

Automatic translation

Logos executes its draft translation independently of the translator. In typical operating mode, the translator selects the text (or series of texts) to be translated from the word processing library. The ensuing Logos translation is fully automatic.

The program returns each completed translation to the word processing text library for subsequent post-editing by the translator. The Logos translation can be considered as the first draft or as a quick scan for information. However, while Logos is broadly capable, and never forgets a transfer, it is not a skilled, professional human translator. Before a translated document is ready for publication and distribution, a competent translator must review it, and edit it using the word processing powers of the Wang system. The human translator is essentially the manager of this process - not a limiting variable.

Multi-media input

Being resident as an integrated function on this leading word processing system instantly provides Logos with all the capabilities associated with that system. Logos commands operate through standard word processing keyboards. Optical character readers, local network and long line telecommunications capabilities and special printers for output, are all standard attachments to the Logos/Wang systems. The integration of natural language translation with word processing is a powerful and natural combination of functions.

Format fidelity

To someone unfamiliar with translation, the format of a document may seem to be relatively unimportant when compared with the difficulties associated with producing an acceptable translation. The practical experience of Logos has proven this not to be true. In nearly every case translations need to be format-faithful to the source to be acceptable. Once again, Logos's integration into the word processing system has a significant advantage. The Wang WP system is already successfully interfaced to a wide variety of other manufacturers' systems and can maintain format control acceptable to the Logos translation system. Format integrity is neither simple nor easily achieved when one is transferring text from one system to another. It is further complicated when the systems are of different manufacture and when telecommunications are involved. Logos and Wang have successfully dealt with these issues.

Semantic power

Logos is semantically strong; this strength is at the heart of our system, and is its measure.

The Logos system can render different context-dependent translations for the same source word, depending on that word's specific use in each sentence, especially if it is a verb (or a word with a verb-derived or verb-related sense).

This is one of the most powerful of the many Logos product features, and it is worth briefly looking inside the system - at what we call the Semantic Table. This is an ordered set of rules for resolving verbal constituents to their appropriate nuance, and then for transferring this nuance into an appropriate target language transfer. For example, there are rules in the Semantic Table handling over ninety senses of the verb 'kommen'.

But the practical significance of semantic strength is

simply its ability to get the right transfers. Because the more often we get the transfers right, the less editing time is needed to get the text right.

Multi-target capability

A Logos translation out of one source language can be a translation into more than one target language. Additional programming is not necessary; all that is necessary is to have a multi-target data file that defines the additional target languages and appropriate dictionaries.

Learning powers

Logos is capable of accepting new words into its dictionary as soon as they are identified as new. The process of describing new words to the permanent dictionary has been carefully designed to be simple and accurate. Logos's interactive process of dictionary enhancement is called ALEX™: the Automatic Lexicographer. No understanding of programming is needed to make effective use of ALEX. The user has only to answer a few questions put on the screen about the word he wants to add. The exact sequence of questions depends in part on the user's previous responses. ALEX relates the new words to words already in the system via analogy and user query responses. Unique semantic codes are created for the new entries and they are immediately entered, with the word, into the dictionary.

Ability to expand

The ability of a very large logically complicated system to remain flexible and expandable has increased significantly in recent years. The highly visible improvements in the price/performance of computer hardware have largely overshadowed the tremendous progress made in software architecture and design. Systems conceived and implemented fifteen or even ten years ago bear little resemblance to similar systems being implemented today. Logos makes extensive use of tables, a highly efficient database approach to dictionary management, and an operating system that enables efficient program execution on a standard word processing system. I might add that Logos's development effort utilises a mainframe whose memory has a capacity of 8 megabytes. It was no small accomplishment for the company's programming staff to convert Logos software to run in less than 1 per cent of the space of the original - that is the 64 kilobytes of the Wang microprocessor. The conversion did not produce any degradation either in quality of translation or in any other feature of the system.

Full-range translation

It is the stated objective of Logos to translate the full range of informative material that is the appropriate source for MT. The present German-to-English general dictionary contains over 100,000 entries. In addition, there is provision for a large number of subject-matter-coded sub-dictionaries, for company-specific dictionaries, and for private-confidential dictionaries. The Logos approach to semantics will enable the system to continually improve its ability to translate complex non-technical material.

Standard, yet customer-specific system

The concept of a customised 'just for me' system may have initial appeal to the first-time user, but practical experience would quickly convince otherwise. A custom system requires custom support, and custom support is at best expensive, and at worst not always available when needed, price notwithstanding. On the other hand, a user has a right to expect his system to be responsive to his particular needs and as such <u>act</u> as if it is custom. The question is how to manage to be standard and yet appear to be customer-specific. Logos has done exactly that. The system consists of a core containing the translation code and appropriate interfaces. The Logos dictionary is a separate file to which the customer can add new words as well as new words supplied by Logos. Software releases for dictionary and linguistics improve the system without disturbing a customer's previous additions.

Cost-effectiveness

Determining the cost-effectiveness of any computer application can be very complicated. But put simply, the translation system must have a cost advantage over the existing system. Computing this advantage is complex and requires the inclusion of all relevant costs. However, in most cases companies only look at the direct cost per line of internal and external translation.

But translation can have an impact on a business far beyond the size of its operating budget. Consider that translation may be keeping a new product out of its foreign markets because of backlogged documentation. The inaccurate translation of a patent could cost a company millions; its accurate translation could save millions. The benefit of a general-purpose translation capability that is timely, accurate, flexible, and instantly available can be of great value to a multinational organisation. Logos on word processing offers such a capability.

Easy to use

The terms 'friendly' and 'easy to use' are overused. I will stand on the widely recognised reputation of the host computers, the Wang OIS or VS, as being easy and efficient to use. Further I can assure you that all Logos screen functions have been carefully designed to be in harmony with the Wang command screens. We believe that Logos is exceptionally user-friendly.

THE ROLE OF THE TRANSLATOR

In the eleven-point description I have just given, I have several times referred to Logos's user - the translator. The translator who works with the system

- initiates and schedules translations;
- post-edits machine output;
- interacts with the Logos dictionary.

Translators schedule translations during the work-day, and Logos does them working day and night if necessary. The computer system can be hard at work translating in the background while other functions such as text entry, post-editing, new word entry or translation management are actively going on at each or any terminal.

And a Logos system can have thirty or more workstations. Text to be translated can be entered at any WP station and edited at stations in the translation department. What has really happened is that the translation function <u>and</u> the translator have been brought into the mainstream of the organisation. Translation is no longer isolated. We believe that Logos increases the productivity of translators and thus their value to the company. Further, we believe that there is a large <u>latent</u> market for translation: translations that just aren't done at present because of cost, or translation backlog. How many copies were made when carbon paper had to be used? Xerographic technology dramatically increased duplication activity because it became easy to do. We believe Logos technology will do the same for translation, and position translators in the mainstream of the business.

SUMMARY

It is impossible to compress literally hundreds of man-years of effort into twenty minutes, and this paper gives only a glimpse of what we believe is the most advanced commercial

language translation system in the world. As the new arrival on the machine translation scene we expect to be examined critically, and hopefully with interest.

AUTHOR

Ralph E. Hawes, Vice-President, Marketing, Logos Corporation, 100 Fifth Avenue, Waltham, MA 02154, USA.