LOGOS
BEATS THE WHEELS OF
FORTUNE

by Andrew Joscelyne

Logos deserves to succeed, if for no other reason than the persistence they've displayed in the face of outrageous adversity. But they also happen to have one of the best batch machine translators in the business. And now, with their new pre- and postediting workstation Filius, based on the much praised Nota-Bene wordprocessor, has Logos collected all the pieces necessary to become king of the machine translation hill? Ace reporter Andrew Joscelyne went to Frankfurt to find out.

TRIALS AND TRIBULATIONS

Imagine you've got a great idea for using computer power to handle translation automatically. You dredge up the necessary capital and found a company to make perfect a product you really believe in. Then, in line with the geopolitics of the day, you opt for a language pair geared to the strategic interests of your country's government. A clear winner, right? Wrong! Whereas automatic translation systems stick around, the ways of the world change. And unless you can get somebody interested, you're left holding your perfect, but unwanted, baby.

One company that knows this story inside out is the Logos Corporation, producer of Logos automatic

translation software. Not once, not twice, but four times over the past fifteen years, Logos has assembled machine translation packages for interested clients — only to feel the winds of political or commercial change blow the wrong way at the wrong moment, and see itself get left in the lurch.

It all began in 1969, when company founder and current Chief Scientist Bud Scott decided to marry his ideas about computers to his knowledge of Vietnamese, picked up while doing intelligence work in the U.S. Air Force. "I had this idea in my head of how to do machine translation," he recalls with the simplicity of conviction. "Of course we experienced a lot of skepticism from other people back then. But I figured we had a sound approach."

A true believer. His Vietnamese translation rig was ready at the right time. By 1970, the war in South East Asia had reached the "hearts and minds" stage. Everything had to be translated into Vietnamese. And Logos Mk. I, by that time fully operational, was called in to break the back of technical documentation in all branches of the armed services and industry. The only problem was that in the end the other side won. South Vietnam crumbled. And Ho Chi Minh City didn't need US helicopter manuals.

Then, in 1973, the CIA thought they could use an English-Russian pair. But it was just a phase they

were going through, and they pulled out on the eve of the Nixon resignation, despite the fact that Logos managed to make them a working prototype. Anybody want to buy an automatic English-Russian translation module?

No thanks. But on the other side of the world, the Shah of Shahs, Iran's Reza Pahlevi was in the market for US military equipment. And no doubt His Magnificence's suppliers had told him of Logos's sterling work in Vietnam. So in 1979, Logos delivered a demonstrable prototype English-Farsi module to the Iranian government, designed to handle both technical and commercial documentation. Unfortunately, by that time, the Shah was far too busy stuffing his suitcase with jewels to care much about Logos. And although the company was finally paid for the job years later, the software, much like its Vietnamese counterpart, was left to turn a magnetic shade of yellow on some dusty Fundamentalist shelf.

Three very special language pairs on your hands, and not a client in sight. Your backseat MT pundit might suggest that the moral of this cautionary tale is "Stick to West European languages." Well . . . in 1980, Logos got a call from Siemens. The West German electronics giant was about to clinch a multibillion dollar American telecoms deal that promised documentation work for the equivalent of 200 translators. To reduce the costs, it wanted Logos to develop a new German-English pair. But just as Logos delivered the goods . . . Siemens lost the contract. And it was Logos to Log-off overnight.

MORE PAIRS

"You pay your dues to sing the blues," goes the adage. And at Logos, don't they know it! But through all these trials and tribulations — though no one got rich — the company certainly chalked up some invaluable MT experience. Most crucially, it built up two high-quality source language templates — one for English and the other for German. This was a major asset. In any language pair, the source analyses gobble up by far the lion's share of the work, time and money. "It's up to five times

When the proposed English-Italian module is complete by mid-1989, you will be able to instruct Logos to translate your input text into any one or all of the four targets at one sitting.

more expensive to do a source than a target," says Scott. "We can now do a target in a year. But a new source may take anything from three to five."

So, armed with its brand new German source, in 1981, Logos opened a European subsidiary: Logos Computer Systems Deutschland GmbH. With a European corporate president and managing director, the company started looking for European customers. With its U.S. head office in Mount Arlington, where the research and technology section still operates today, the German outfit set up shop in Frankfurt, the nerve center of an export-driven industrial zone with a major airport and serried ranks of hightech multinationals at its heart.

After four years marred by the kind of strategic mistakes immature companies are prone to, Logos finally took off in Europe in 1985. By that year, it was offering more market-sensitive language pairs to ambitious German exporters than just the original German to English. And with major high-tech clients like Nixdorf, IBM and Hewlett-Packard – in both Grenoble and Stuttgart – along with automaker BMW, and others, it was looking like Bud Scott's "sound approach" was finally paying off.

The language pool today includes not only the German-English grandaddy, but English-French, English-German, and – since late 1987 – German-French. Meanwhile, English-Spanish is on the testbench in the US, notably at Unisys, for which Logos translates technical manuals. And rumors suggest that this module will be available in Europe by the late summer.

Until very recently, Logos generated each language pair as a single module, with the French in German-French quite independent of its counterpart in English-French. These days, though, the system takes a multi-targeting approach, in which the single source analysis – taking up some 80% of the program – will convert to any one of a range of targets. When the proposed English-Italian module is complete by mid-1989, you will be able to instruct Logos to translate your input text into any one or all of the four targets at one sitting.

BIG CLIENTS
Logos was originally developed on IBM, and runs today under CMS and MVS operating systems. When the company first started, the best wordprocessing technology was from Wang. And since Fred Wang himself was interested in Logos's work, he commissioned the company to port its software to a Wang VS compatible workstation. Years later, it was chuckles all round when Logos personnel visited their client Nixdorf, to see that the

computer maker had installed a rival Wang machine dedicated to translation work. Soon Logos will be available under Unix, first on the Sperry Univax 5000, and then portable to other environments.

There is an equally intimate Big Blue connection. IBM is not exactly paying for translation work done, but Logos is installed in the IBM Computer Center in Hamburg and – since 1986 – in Zurich too, making it available on a time-share basis on the IBM network in both Germany and Switzerland. "This is particularly useful for us," says Philip King, pre-sales support person in Frankfurt, "since foreign companies find it very hard to comply with Swiss law when they want to open an office in the country. And we are developing an interesting clientele in the banking and insurance business there".

In another Big Blue move, Logos has been installed at the IBM European Language Support Center in Copenhagen, where it's being independently tested for productivity. And another recent Scandinavian client is the Swedish PTT, Televerket, which recently needed 4000 pages' worth of technical manuals written for the German BTX videotex system translated into English. Logos duly obliged.

THE USER SCENARIO
A typical user scenario goes like this. The documents to be translated are entered via a wordprocesser on an IBM or compatible PC, converted into host-readable format using 3270 emulation, and transmitted via modem to the host computer. Here, Logos can work fully automatically in batch mode, throughputting some 3000 lines (around 100 pages of A4) per hour – or 4 to 5 pages an hour on the Wang.

Since the work can be done overnight – *nachts, wenn der Übersetzer schläft* goes the soothing slogan – it is returned to the user in WP format, allowing postediting work to be done on the inhouse PC. Online time spent logging on, sending off text, and next-day downloading onto your PC for postediting can be counted in seconds, thanks to a dedicated button that avoids the "hot-key" routine.

The idea behind the Logos strategy is to let the automatic system do the dreary repetitive work it is particularly good at, thereby freeing skilled professionals to handle advertising, commercial or legal texts, which demand the kind of style-sensitive intelligence that software doesn't possess. Machines putting translators out of work? "Not a bit of it," says Philip King. "One company that decided to use Logos eventually took on three *more* translators when it realized how much more work could be done with it. Logos actually creates jobs by allowing companies to translate everything they would like, not just the minimum they must. We are concerned with generating productivity, not dole queues."

For translators who work with Logos, however, there is the inevitable teething period as they adapt themselves to the technology. This may mean switching from dictaphones and pencils to WP keyboards – "There are plenty of inhouse translators in Europe who still think that a mechanical typewriter is high-tech," smiles King. Or it may mean learning how to relate to the two current translator aids that come with the Logos package.

First of all, the client has to build up his own lexical database. And this can take up to three months. Each term has to be coded, first with subject matter codes, of which there are hundreds, running from Archeology through Metallurgy to Zoology. Then each term has to be given company-specific codes for access. The word "plant," for example, is coded under "Industrial Organization" for certain clients and "Agriculture" for others.

Translation agencies that offer multi-field translation services are more or less excluded as potential Logos clients on this count – they'd be spending all their time building dictionaries for the many subjects they'd want to offer.

The basic installation termbase arrives with subject-specific dictionaries of some 15,000 words, as well as the general language dictionary of up to 145,000. However, as King says, "We don't know what the client actually wants to say. So from his point of view, the dictionaries might as well be empty." Even though many current clients use Logos to handle the translation of technical documentation in the dataprocessing field, there is still sufficient variation in company terminology policy for the large and varied dataprocessing dictionary to need customizing.

The dictionary update tool used to help the translator build that company-specific list is called *Alex*, which uses an interactive query process to enter such information as subject matter codes and the part of speech of each term. When the input text has been loaded onto the translation system, the first thing Logos does is perform a New Word Search. The output from this search becomes the raw material for the Alex-wielding translator to work upon. Most users find the facility straightforward to learn and, inevitably, less and less necessary as company terminology becomes uniformized.

A second, more sophisticated, tool available on Logos ic called *Semantha*. This allows the translator to actually construct customized semantic rules to handle the different meanings of, for example, "make" in collocations such as "make a mistake" and "make a cake," by means of generic linguistic concepts. The rules then merge the semantic info in the dictionaries with the syntactic analysis of a sentence, generating an adequate translation of special phraseology or idioms.

At Hewlett-Packard in Grenoble, a translator testing Logos's English-French translation of dataprocessing manuals, found that Semantha filled a necessary gap in the system, and improved output. "Sentences came out badly translated from the current rules. But entering updates with Semantha was both practical and efficient. Though you had to make sure that your new rules didn't conflict with other rules already in the system".

FILIUS
If Logos sounds pleased with itself today, it's not because it has achieved a completed viable system. Logos is a "living product" – says its maker – and gains at least one new release per language pair per year, yielding – say its users – significant differences to productivity. Especially those of the 1987 vintage. "It's very important to specify which version you're working with when you make an assessment of Logos," says the Hewlett-Packard translator. English-French Release 4.1 was ready in January of this year, and is still under assessment.

No, the real reason why Logos can't wipe the smile off its face is more domestic. According to the joke doing the rounds at the company, *Semantha* and *Alex*, that wordy couple, have – albeit incestuously – begotten a son, suitably named *Filius*, who was recently presented in Europe at the giant Hanover Trade Fair. And according to first reactions, he promises to be the best thing to hit translators since bilingual dictionaries.

When you're trying to sell machine translation as a productivity tool, there are two approaches you can take. One is to vaunt the system itself – and sink your updating resources into readjusting those hundreds of thousands of lines of machine code to

When you're trying to sell machine translation as a productivity tool, you can take two approaches. One is to vaunt the system itself – and sink resources into readjusting hundreds of thousands of lines of machine code to improve raw translation quality. The other is to study the ergonomics of the workstation – to help the translator work better. Logos tries to take both approaches in two different ways.

improve raw translation quality. The other is to study the ergonomics of the workstation – to help the online translator work better. Since the current translation market is, in Bud Scott's words, "muddy," with users making different demands at different points in the process, Logos tries to take both approaches in two different ways.

On the one hand, system updates will generate improvements in Logos's effectiveness as what Bud Scott calls "an information utility, aimed at aiding management to communicate across linguistic borders," without trying to reach postedited quality. Scott figures that this use for automatically translated raw output marks "a new era" in the history of technology.

On the other hand, there's an immediate need to increase the translator's daily pre- and postediting output. And this is where young Filius will come into his own. As one Logos tester has suggested: "If the translator can concentrate solely on terminology and postediting, then Logos's advantages as a wordcruncher with an infallible lexical memory are obvious." Of course, the extent to which preediting of the source can accelerate the whole production process varies from case to case. But Logos figures that this front-end work can be greatly enhanced by the right kind of interface technology.

DRAGONFLY

What Logos has done is to incorporate a value-added edition of version 3 of super-WP *NotaBene* (see LT#6 for review), from Dragonfly Software of New York, into a customized translator's workstation – a "translator's productivity tool" – which can either stand alone as a PC facility or, ideally, function as a front end to the Logos mainframe system, including the Alex and Semantha updating applications. The benefits of Filius to the working translator have been boggling the Logos staff themselves, let alone the previewers who've so far had a look.

To begin with, there are the built-in speed and extensive document editing facilities of the original NotaBene, already much praised in the dedicated press. The real translator boons, however, lie in the split-screen option with its accompanying focus command that will highlight source and target sentences at the same time. Plus – and this is a real timesaver – a simultaneous scroller that shifts the source text as you move your editing cursor up or down the target material.

Column and cellular table controls allow you to recreate the same tabular configurations in the target as in the source, whatever the length differences between the two. If, on the other hand, your document has intext tables that can't be translated by the system due to formatting differences, Filius allows you to protect them from Logos by opening a window-file for the tables, and then either translate them apart or copy them from the window and reinsert them in their original form in the target file.

Filius is particularly adept at dealing with typos, wordblindness and people with holes in their memories. For the practicing translator who makes typing errors, the first way of getting a correction is to send his document off to Logos itself and ask for a new word search that will automatically send back a list of mistaken words.

On Filius, though, you have a spellchecker that posts up phonetically similar forms, from which you can select the one you want. If, however, you are forever committing those customized errors like "recieve," "catagory" or "informtaion", you can get Filius to enter a list of them into a Personal Spell dictionary, which can then be loaded when you write and automatically correct your typos while you work. In addition, there is a last-two-letter, word, or sentence transposition key for speedy switch of elements.

So far as dictionary handling is concerned, by using the window opening facility, the user can have multiple wordlists onscreen cohabiting with the two translation texts. At present, the built-in dictionary and thesaurus are only available in English, but the environment is designed to let the user build up his own dictionary. And you can call up the hyphenation dictionary and enter your hyphenation rules for own-language terms. For example, when you use the column facility, Filius will automatically hyphenate your words according to your rules.

REDLINING

Judging by Dragonfly's general willingness to produce any update a customer might want within three weeks – they will customize it for any printer in the universe that exists in more than one copy, says Dragonfly's blurb – hyphenation dictionaries will no doubt eventually be available for any idiom you care to imagine, except Basic perhaps. To profit from the system's configuration, it is best to create lots of small dictionaries with a maximum of 15K of words. The built-in thesaurus is read only, which makes sense for Logos since company-specific glossaries can be held in the personal dictionaries.

From the language point of view, Filius comes with West European keyboards plus dead accent facilities for adding diacritics to the basic US keyboard configuration. You can also use Swedish

characters. An additional language supplement demonstrated recently at Hanover offers three columns of text handling three different alphabets, with the Hebrew characters unfurling right to left at the flip of a key. For further onscreen editing, redlining and bluepencilling facilities allow you to run through a text, and delete terms still coming out wrong from the translation system. You can collect them into a file while you're still editing by opening a new window, and then send them off to the Logos dictionaries for updating.

A final, but in Logos's eyes crucial feature of Filius, is the oft-praised NotaBene textbase manager, which allows you to search and retrieve any bit of any text however long from anywhere in your files. The significance of this for professional translators who work in companies with hefty documentation needs, can be illustrated with a real-life example. Eppendorf Gerätebau GmgH is a Hamburg-based company that produces medical appliances needing extensive technical documentation. To handle the English translation they have used an inhouse Logos system for the past four years and are more than happy with it.

One of Eppendorf's specific translation tasks is updating and then retranslating portions of already translated text. When and if Eppendorf invests in Filius, it will be able to do this using a textbase feature that recognizes unchanged sentences and sends only revised text to Logos for translation. Smart, *nicht wahr*? "Since Logos provides software for automatic translation," says Director of Marketing John Cave, "it was a natural extension for us to address the ancillary activities of document editing and text management."

POSTES CANADA POST

If all this isn't enough to make 1988 a year for Logos to forget its commercially bumpy start in life, then news that a major potential client, the Department of the Secretary of State of Canada, has decided to renew credits to continue testing Logos into a second year of English-French translating ought to do the trick. In the first phase, Canadian officials had anticipated a productivity increase of from 20 to 30%, but after five months they already noticed a 100% increase, obtained without degradation of translation quality.

A second version upgrade was introduced in July 1987, with a correction rate of 36% of the errors of the first version. The third beta version is now installed and the number of sites will increase from one to three and then to five, serving various Canadian government departments and agencies. For Logos, the real bonanza would come if the Canadian Government decided to translate French documents into English, since this would entail building a costly new French source analyser.

Says Bud Scott. "If the Canadian government were to finance us, we would be interested. But we're also thinking about Italian as a possible source." In the dead of night, though, the ghosts of crises past loom up, and Logos staff must be secretly hoping that the Canadian government doesn't suddenly renege on its bilingual policy, or declare Hopi the national language.

Andrew Joscelyne *is LT's Paris-based senior contributing editor.*

Logos Corp., *1 Dedham Place, Dedham, MA 02026, U.S.A., Tel: +1 (617) 326 7600*
Logos Computer Systems Deutschland GmbH, *Lyonerstr. 11, 6000 Frankfurt/M. 71, F.R.G., Tel: +49 (069) 666 6950*