[*Machine Translation Summit*, September 17-19, 1987, Hakone Prince Hotel, Japan]

THE METAL SYSTEM. STATUS 1987

Dr.Thomas Schneider
Siemens AG

1. History
2. Hardware
3. Grammar
4. Lexicon
5. Development Tools
6. Current Applications and Quality
7. Research, Future Applications

1. History

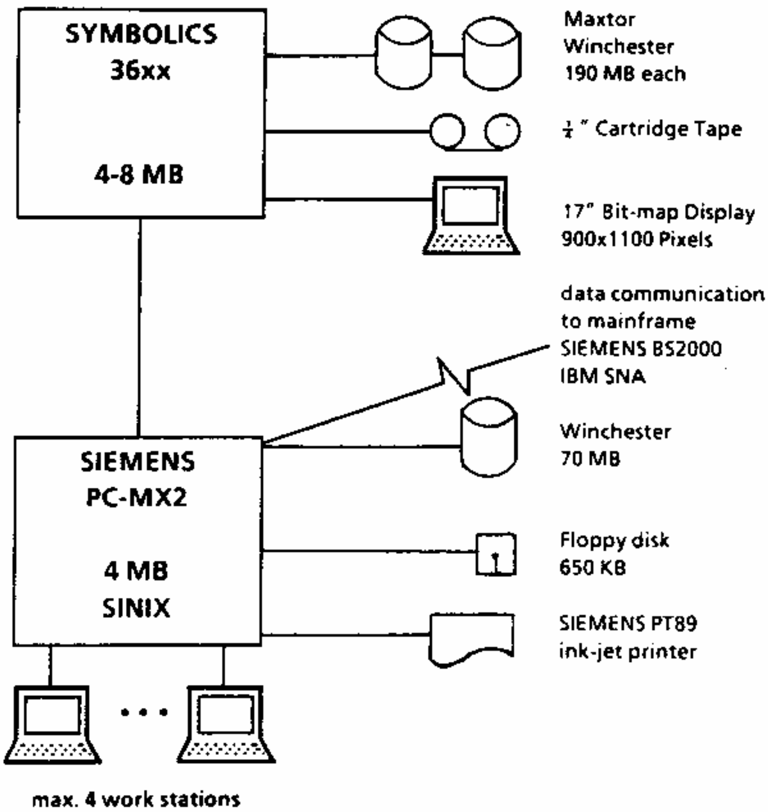
In the late seventies, when there was a noticeable shortage of qualified technical translators versus the volume of required in-house translations, Siemens began to look for an operative machine translation system. It was intended to increase the productivity of the translators available, and to reduce the time required for the translation process. This is extremely critical if voluminous product documentation needs to be delivered on time.

After experiments with some of the systems which were offered commercially, it seemed that for our purposes a broader linguistic analysis was required. So in 1978, Siemens entered into a cooperative agreement with the University of Texas at Austin to develop the METAL system which at that time existed as a somewhat unwieldy prototype but which was based on more than twenty years of theoretical research at the university.
Whereas at first the project was expected to run for about fifteen to twenty years, keeping in mind all the necessary research, progress came more rapidly than anticipated. Today the German-English version is being used productively in several pilot operations, and work on other language pairs is approaching a level where the first tests can be conducted.

2. Hardware

At present, METAL is implemented in CommonLisp and runs on a Symbolics 36-series LISP-machine. Since these LISP-machines are single-user stations but the system's throughput far exceeds the amount of text a single translator could postedit, the LISP-machine has been linked to a multi-user PC, the Siemens PC—MX2. For considerations of response time, it is recommended that not more than three posteditors work simultaneously even though it is physically possible. METAL can run in batch in the background while formatting, postediting and the administration of translation tasks are

handled on the PC. For reasons of lexicon integrity, the functions of lexicon modification and structuring reside on the "central" LISP-machine.
The hardware configuration looks as follows:



< figure 1 >

Of course, there is a world of difference between a prototype that can translate individual sentences from the keyboard and a productive system. Most of the texts which need to be translated quickly and are of great volume such as e.g. technical documentation are heavily formatted. In some texts, more than half of the characters may be non-translatable material, e.g. flow charts, diagrams, tables and control characters for layout and print. It would be prohibitively expensive to manually extract the text portions to be translated and afterwards manually reinput them. Besides the expense, it would also be another source of errors.

Therefore, METAL was designed as part of a chain of processes, from text acquisition via deformatting, spelling correction via translation to automated reformatting procedures.
A text is usually received in machine-readable form (usually written on a word processor or else read in via OCR). Several programs running on the PC-MX2 check the pages for tables, graphs etc and mark them. They identify the text portions to be translated and generate a mask of the page. The individual translation units, usually sentences but in the case of headlines or table entries also noun phrases or single words, are automatically recognized,

numbered consecutively and extracted from the page mask. They are written into a File and transferred to the LISP-machine for translation.
The file of the translated units is returned to the PC for postediting.
Here, the posteditor can choose whether he wants to postedit an interlinear version that groups together individual source language/target language equivalents sentence by sentence, or whether he prefers a target language output that has already been reformatted.
In the former case, the posteditor would start the reformatting after having made his corrections. After this step, the target language text is available with all the formatting commands and with the same layout as the original.

3. Grammar

METAL is designed in a highly modular way so as to permit the inclusion of new elements or the modification of existing elements without any ill effect on the other components. As there is at present no linguistic theory available that would describe even a single language unambiguously a somewhat eclectic approach has to be chosen.
METAL employs a transfer approach rather than an interlingua. It seemed that to define a meta-language incorporating all possible features of many languages would not only be an endless task but rather fruitless as well. The system would soon become unmanageable and perhaps collapse under its own weight. If, however, the meta-language were reduced to a manageable level of abstraction then too much information necessary for a faithful translation would be lost. Tests with several European languages have shown that at least between these related languages a transfer system is probably adequate.
Actually, the identical analysis of a source language seems sufficient for simultaneous translation into several target languages.

METAL uses phrase-structure rules which are augmented by tests on the individual constituents and various other constraints. As the rules are recursively applied, in a Markovian manner, their number can be kept low.
The grammar rules are indexed to make processing more efficient.
The most commonly applied rules, e.g. those for word level morphology and for frequently occurring basic structures, are defined as the most basic level. Higher level rules deal with more complex or even ungrammatical structures. If a structure can be interpreted using lower-level rules then the more complex and less likely rules are disregarded which saves processing time. If no interpretation is possible with the lower level rules then incrementally higher levels of rules are added to the lower level rules to attempt an interpretation.

METAL uses a bottom-up chart parser. Unlikely paths are eliminated via preferential weightings calculated from lexical and grammatical data.
This procedure not only increases processing speed but usually provides the best interpretation for transfer to the target language.
If no plausible interpretation of the complete sentence is reached, the system will go into a fail-soft routine and output a translation of the individual phrases it has been able to interpret.
Surprisingly often, the output is still a grammatically correct translation of the original sentence, at least going from German to English; in other cases, the posteditor has to correct the output.

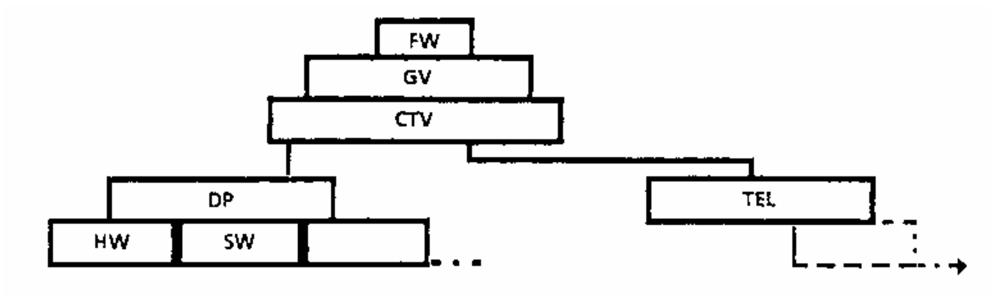
4. Lexicon

METAL operates on monolingual lexicons and one transfer lexicon for each language pair. The monolingual lexicons contain morphological, syntactic and semantic information needed for the analysis and/or the generation of a language. The transfer lexicon provides a link from the source language to the target language, indicating in which contextual environment and in which subject field the source language entry should point to a certain target language entry.
The advantages of such a structure are obvious. The extensive information contained in the monolingual lexicon needs to be carried only once, even if many different entries in one of the languages correspond to the same entry in the other language.

For verbs, the monolingual entries contain information about inflectional patterns, legal frames and - if necessary - further restrictions on the arguments (e.g. case, semantic types etc). Different stems are subsumed under the canonical form.
Nouns are specified as to inflectional class, grammatical gender, natural gender, occurrence of determiners, specification of complements and semantic type.
Adjectives are coded for inflectional class, possible grammatical functions. position in regard to modificand, adverb derivation, comparative and superlative derivation and semantic type. Similar information is added to adverb entries.
Whereas it is possible for an end user to also modify entries for function words such as conjunctions, prepositions and determiners, it is not recommended since these structurally significant words are closely linked to the grammar rules.
The transfer lexicon permits significant syntactic transformations. On top of being able to specify transfers on the basis of the instantiation of frames, the presence of arguments of a certain type or of a specific canonical form, the user can change the target structure considerably.
Surface structure active phrases can become impersonal constructions, roles of arguments can be changed, complements can be converted, elements can be deleted or added etc.
All these options are available to the end user via a menu-driven INTERCODER which has proven its usefulness in reducing coding time by a factor of ten. The INTERCODER is reasonably, intelligent as it guesses at the necessary coding of entries provided it is given the canonical form and word class. The missing pieces of information are inferred from information already contained in the lexicon and from a set of heuristic rules.

To build a productively useful MT system it is imperative to include a sizeable dictionary. However, an unstructured voluminous dictionary can create more problems in regard to unwanted ambiguities than would be gained by having extended text coverage of the lexicon.
On the contrary, among European languages a set of about 5000 words make up about 90 7. of the average text but beyond that the point of diminishing returns is soon reached. Moreover, MT systems are not intended to translate general language or even literary texts.
MT is aimed at texts which are of sufficient volume and have to be translated within a very short time. In such texts, e.g. product documentation, the percentage of general language vocabulary is quite limited while subject-specific terminology abounds.
This makes it possible to modularize the dictionary and assign

preferences for specific transfers based on the subject area of the text to be translated.
The METAL lexicon is organized as follows: There are modules for function words (FW) like prepositions, articles and conjunctions, for general vocabulary (GV) and for common technical vocabulary (CTV) organized in a tiered hierarchy. On the next level down, there are subject-specific modules which in turn can have subsets of more specific terminology. The module "Data Processing", for example, has subsets called "Hardware", "Software", "Data Transmission" etc. Should any further specification be required far e.g. customer—specific or product-specific terminology, new modules can be added to the structure. There is also the possibility to define a transfer on the basis of "dialect". Thus a text intended for Great Britain will automatically show "boot" instead of the American "trunk".
The METAL lexicon structure can be visualized like this (simplified):



< figure 2 >

Before a translation run is started, the modules appropriate to the subject area of the text are defined. If the syntactic and semantic criteria for the selection of a lexicon entry are met and there are competing transfers then the one tagged for the subject area of the text or tagged for a hierarchically closer module is chosen.

The main source for the required terminology for new subject fields is TEAM, the multilingual terminology data bank operated by Siemens which at present contains ca. 2 million records in up to eight languages. An interface between TEAM and METAL facilitates the installation of new lexicon modules. But in productive use, the end user is called upon to update the system dictionaries as the need arises.

5. Development Tools

In the course of the METAL development, several tools have been written to aid in lexicon and grammar work. For lexicon maintenance, there are utilities to check for syntactic errors in individual entries, and others to ensure the integrity of the data base after changes have been made to some entries. General statistics modules are available as well.

For grammar development, it can be useful to have statistics on the frequency of application of specific rules and information about the number of cases in which a specific rule contributed to the parse of sentences. Another program permits the graphic display of analysis trees or generation trees. It has proven to be a significant aid in spotting errors, especially since all information carried by each node can be made visible at the push of a key, and the rule responsible for the construction can be edited on the spot.

A set of benchmark texts is used to point out differences in translation quality from one version to the next. A difference list can then be used for diagnosis. Also for diagnostic purposes is the step-by-step reconstruction of attempted rule applications and their effects.

These tools, however, are intended for the system developers only and are not made available to end-users.

6. Current Applications and Quality

So far the German-English version of METAL has been used for the translation of several thousand pages of different texts in the subject areas of data processing, telecommunications and related areas. Besides our in-house pilot installations, there are now external customers working productively with the system.

Translation speed is at present about one word per second which amounts to more than 200 pages per day.

Sometimes the quality of machine translation is expressed in percentage points. But as with human translation, it is almost impossible to quantify the quality of machine translated output. It is at best questionable to operate on the basis of a percentage of "correct" words since that may depend entirely on the type of text, the author's personal style as well as other factors. It is also difficult to decide if a "correctly" translated sentence has to be polished stylistically or not since that depends on the purpose of the text, the expectations of the intended readers and rather idiosyncratic likes and dislikes.

Test with METAL have shown that usually more than half of all sentences were translated "correctly", and that the rest was at least good enough to serve as a basis for postediting. More important, however, is the fact that translators (not just biased developers) considered the output acceptable enough to volunteer to use the system in their daily work. Spot checks showed an increase in productivity by up to a factor of eight but this is an unrealistic target. It takes too much concentration to postedit large quantities of text and check for errors of content which may have been introduced in the source language original. Nevertheless, based on extensive pilot applications the hope for an increase in productivity by a factor of three does seem realistic.

7. Research, Future Applications

At present, METAL is used productively for machine translation from German to English. Work on English as a source language and Spanish as a target language is under way, and prototypes should be available soon. By the end of the year, components for Dutch and French can be expected as well.

Even if among European languages syntax-based approaches lead to useful results more work needs to be done on the semantic component. One of the major problems in this field is that there is no solid theoretical base for a language-independent semantic representation that would lend itself to the description of real-life texts. Needless to say, a machine translation system has to be able to translate texts as they are written - and not as they should have been written. A competence model in a miniature domain may serve for experimental purposes but there is no certainty that such an approach could ever be extended to a performance-oriented application in larger subject areas.

At present, a research project in cooperation with the University of Munich seeks to identify possible venues to a semantic description of lexical items and their interrelation with a view to disambiguate structures that cannot be resolved on a syntactic (plus semantic feature) level. It is not yet clear if an adequate and manageable solution can be found. Most promising at the moment seems a combination of several approaches including prototype semantics, collocational statistics and taxonomic hierarchies including synonym references.
Even further removed is the question if the chosen approach can also be used for information retrieval on the basis of semantic content - which would extend METAL's scope to many other applications beyond machine translation.