Jonathan Slocum
Linguistics Research Center
The University of Texas

INTRODUCTION

Although funding for Machine Translation (MT) research virtually ended in the U.S. with the release of the ALPAC report [1] in 1966, there has been a continuing interest in this field. Rapid evolution of science and technology, coupled with increased world-wide exposure of their products, demands more and more speed in translation (e.g., in the case of operation and maintenance manuals). Unfortunately, this rapid evolution has made translation an even more difficult and time-consuming task. The large surplus of (presumably qualified) translators cited by the ALPAC report simply does not exist in many technical areas; the current state of affairs finds instead a critical shortage. In addition, the proportion of scientific and technical literature published in English is diminishing. As qualified human translators become more scarce and costs of human translation rise while costs of purchase and operation of powerful computer systems fall, there must come a time when, if MT is feasible at all, it will be cost-effective. It is appropriate, then, to investigate the state-of-the-art in MT with respect to two central questions: is high-quality MT feasible (and in what sense); and if feasible, is it cost-effective?

This paper reports the results of an experiment in highly automatic, high-quality machine translation. The LRC's MT system, METAL (for Mechanical Translation and Analysis of Languages), is an advanced, 'third generation' system incorporating proven Natural Language Processing (NLP) techniques, both syntactic and semantic, and stands at the forefront of the MT research frontier. In the experiment, METAL was employed in the translation of a 50-page text from German into English in order to determine whether the system as it exists can be effectively applied to current translation needs, effectiveness to be determined by some objective measure of the quality and cost of machine (i.e., METAL) vs. human translation.

EARLIER MT EFFORTS

Since Bruderer [2] has recently published a complete survey of MT projects, and Hutchins [3] reviews the most important developments through 1977, we will mention only a few of the major efforts. The first popular demonstration of the possibilities in MT was provided by IBM and the Georgetown University group in 1954 [4]. With a vocabulary of about 250 words and a grammar comprising some six rules in what was called an "operational syntax", the system demonstrated some rudimentary capability in Russian to English translation. This instigated a massive government funding effort over the next decade, and some 20 million dollars was invested in 17 different projects. By 1965 the Mark II Russian-English system [5] had been installed at the Foreign Technology Division of the U.S. Air Force at Wright-Patterson AFB, and the Georgetown system had been delivered to the Atomic Energy Commission at Oak Ridge National Laboratory and to EURATOM in Ispra, Italy. Reviewing MT systems such as these at the request of the National Science Foundation, the Automatic Language Processing Advisory Committee (ALPAC) reported in 1966 that MT was slower, less accurate, and more expensive than human translation; further, that there was no predictable prospect of improvement in MT capability. Though strongly and perhaps justifiably criticized [6], this report soon resulted in the virtual elimination of MT funding in the U.S., and a sizeable reduction in foreign efforts as well.

Peter Toma, who was responsible for the installations at Oak Ridge and Ispra cited above, soon began private efforts at improving the Georgetown system. This culminated in SYSTRAN [7], which replaced Mark II at WPAFB in 1970 and the Georgetown system at EURATOM in 1976. SYSTRAN was also used by NASA during the Apollo-Soyuz mission. In 1976 the Commission of European Communities adopted SYSTRAN for English to French translation; however, an evaluation of its translations by the EEC posteditors in Brussels found the results to be far from satisfactory: "all the revisors had exhausted their patience before the end" [8]. Despite its generally low translation quality, SYSTRAN is the most widely used MT system to date. Its chief commercial competitor, LOGOS [9], is another example of a "direct" MT system. As in SYSTRAN, the analysis and synthesis components are separated but the linguistic procedures are designed for a specific source-language (SL) and target-language (TL) pair. In an evaluation by Sinaiko and Klare [10], LOGOS did not fare well. Bruderer [2] reports further development for translation into Russian, and experiments on French, German and Spanish, but provides few details.

In an effort to correct the obvious inadequacies of these and other 'first generation' systems, which essentially translate word-for-word with no attempt at a unified analysis at the sentence level, and which were developed ab initio for a specific SL-TL pair, researchers began to investigate methods of analyzing sentences into structures from which in theory any TL could be generated. There are two broad types of such 'second generation' systems. One type produces analyses in a "neutral" structure, or 'interlingua'; the other produces SL syntactic structures which are transformed via a process called 'transfer' into a syntactic structure for the TL sentence. One example of the former approach is the system produced by the Centre d'Études pour la Traduction Automatique (CETA) at the University of Grenoble [11]. During the period from 1961 to 1971 this group developed a Russian to French MT system. An evaluation at the end of that period revealed that only 42% of the sentences were being correctly translated. Some failures were due to errors in the input, but the majority were due to programming errors, failure to produce a lexical analysis of a word or a syntactic analysis of a sentence, inefficiencies in the parser causing it to apply too many rules, etc. The Traduction Automatique de l'Université de Montréal (TAUM) project [12] is an example of the transfer approach. There are five grammars called "Q-systems" to effect morphological and syntactic analysis of English, then transfer, then syntactic and morphological synthesis of French. Each such stage consists of a series of generalized tree-structure transformations. The significance of TAUM is that, of the second-generation systems, it is the nearest to operational implementation: it is to be applied to the translation of aircraft maintenance manuals.

In 1978 the European project EUROTRA was initiated, apparently adopting the newer Grenoble system ARIANE, in order to produce an advanced, second generation MT system for the eventual replacement of the first generation system (SYSTRAN) currently in use [8]. The Grenoble group, now titled Groupe d'Études pour la Traduction Automatique (GETA), abandoned their earlier approach in light of its deficiencies and produced a system to translate in six passes: morphological analysis, multi-level (syntactic and semantic) analysis, lexical transfer, structural transfer, syntactic generation, and morphological generation. Multi-level analysis, structural transfer, and syntactic generation are all effected via a general tree-to-tree transducer program, some-

what less powerful but perhaps more efficient than the Q-
systems transducer in TAUM; the other components have spe-
cial programs suited to their function. The emphasis in
this project is apparently twofold: increased efficiency
and reliability through adoption of components with the
minimum necessary power, and decreased sensitivity to
failure in individual stages through the expedient of in-
suring that every component has some output, even if
such output is nothing more than the original input. If
we have interpreted the Vauquois mimeo [8] properly, this
must be the largest and most comprehensive MT project yet
undertaken.

DESCRIPTION OF METAL

There are two different classifications of "generations"
in MT systems. The first posits three generations (cur-
rently) according to the following criteria: (1) trans-
lation is word-for-word, with no significant syntactic
analysis; (2) translation proceeds after obtaining a
complete syntactic analysis of an input, with no signifi-
cant semantic analysis; (3) translation proceeds after
obtaining a complete semantic analysis of an input. The
definition of 'third generation' says nothing about ex-
tra-sentential information, and one might posit a
'fourth generation' which employs such information. The
other classification proceeds according to the following
criteria: (1) translation proceeds "directly" from the
SL to the TL, and the SL is analyzed only to the minimum
extent necessary to generate TL equivalents; (2) trans-
lation proceeds "indirectly" by deriving a more-or-less
standard analysis of the input, independent of the TL in-
volved (but not necessarily of the SL), and then genera-
ting TL output based on the standard analysis. Within
this definition of 'second generation', as noted above,
there are the 'transfer' vs. 'interlingua' approaches.
We prefer to characterize METAL as a 'third generation'
system according to the first classification given above
because this makes it clear that METAL derives a sub-
stantial semantic analysis, whereas the second definition
of 'second generation' does not necessarily imply that
semantic analysis of any kind is performed.

METAL comprises two distinct components: the linguistic
and the computational. The linguistic component con-
sists of lexicons, phrase-structure grammar rules, case
frames and transformations. SL and TL lexical entries
include feature-value pairs encoding syntactic and sem-
antic information such as grammatical category, inflec-
tional class, semantic type, and case information (see
Figure 1). Transfer lexical entries indicate how and
under what conditions words or idioms in one language
translate into words or idioms in another (see Figure
2). The phrase-structure rules may be augmented with
procedures to determine their application via feature/
value tests, to add or copy features and values in the
interpretation being constructed, to invoke case-frame
routines, and to invoke specific or general transforma-
tions. Case-frame routines determine semantic case re-
lationships between verbs and nouns on the basis of syn-
tactic and semantic features, and produce their output
in the form of propositional trees. Transformatio - are
pattern-pairs that specify old and new tree structures;
when invoked, a transformation attempts to match its
"old" side against the current structural descriptor,
and if successful converts it into one matching its
"new" side. In the process, features and values may be
tested and set arbitrarily. This provides the grammar·
with virtually unlimited ~ontext sensitivity, but since
no interpretation can affect the operation of the parser
it still enjoys the advantages of context-free opera-
tion. Finally, there is a method for scoring, or rating,
interpretations; this allows the system to determine the
"best" interpretation for translation, and also provides
another mechanism for rejecting the application of any
rule, viz. a score below cutoff. Figure 3 illustrates a
typical grammar rule.

```
(IN                     CAT (PREP)
  ALO (in)     (i)
  GC  (A D)    (D)
  CN  (S)      (M)
  PLC (WI)     (WI NF)     %
  RO  (TMP TOP LOC DST TAR EQU))

(IN                     CAT (PREP)
  ALO (in)
  RO  (DST LOC)
  PO  (PRE)
  ON  (VO))

(INTO                   CAT (PREP)
  ALO (into)
  RO  (DST LOC)
  PO  (PRE)
  ON  (VO))
```

Figure 1
German Preposition "in" and Two
Corresponding English Prepositions

CAT - grammatical category
    PREP - preposition
ALO - allomorph
    'in' - the string "in"
    'i' - (as in the string "im")
GC - grammatical case
    A - accusative
    D - dative
CN - contracted [with]
    S - (as in "ins")
    M - (as in "im")
PLC - placement
    WI - word-initial
    WF - word-final
RO - semantic role
    TMP - temporal
    TOP - topic
    LOC - locative
    DST - destination
    TAR - target
    EQU - equative
PO - position
    PRE - pre-posed
ON - onset sound
    VO - vocalic

```
(INTO     (IN)     PREP  (GC A))
(IN       (IN)     PREP  (GC D))
```

Figure 2
Transfer Entries for
the German Preposition "in"

The German PREPosition "in" (in parentheses) may trans-
late into the English PREPosition "into" if the Gramma-
tical Case of the German PP is 'Accusative'; it may tran-
slate into the English PREPosition "in" if the Grammati-
cal Case of the German PP is 'Dative'. Arbitrary numbers
and types of conditions may be specified in transfer
entries.

The computational component, written in LISP, consists
of the parser, the case-frame routines, the transforma-
tion pattern-matcher, the transfer program, the genera-
tor, and other procedures needed to drive and support
the translation process. The parser is a highly effi-
cient implementation of the Cocke-Kasami-Younger algo-