Machine Translation Project at Musashino Electrical Communication Laboratory

Oct. 1981

Hirosato Nomura

Research Division, Musashino Electrical
Communication Laboratory, N.T.T.,
Musashino-shi, Tokyo, 180 Japan

The machine translation project attempts to build a totally intelligent system which can understand languages and translate to other languages. The system has also the ability to support many tasks relevant to not only translation processes but also other processes including linguistic analysis, knowledge manipulation, programming, etc. We have much interest in understanding the meanings of sentences, the role of knowledge in that process, and man-machine interaction.

The experimental system is now under implementation and expected to operate at next spring. The system is planned to translate from Japanese into English and vice versa. Though the goal of the research is intended to automatic translation, the experimental system has facilities to interact with a human when falls in with unknown words or unusual meaning. Interaction is done via high resolution color display (imagine the Xerox's Star, but ours is advanced to treat KANJI characters with multi window).

The distinguished characteristics are powerful semantic analysis with discource analysis, uniform representation of intermediate structures, dictionaries and knowledge-base, and man-machine interface.

1. Background

a. Artificial Intelligence approach to Machine Translation.
b. Semantics and understanding is a main theme.
c. Man-Machine interaction is another theme aimed to support
    linguistic analysis,
    programming,
    linguistic data (including dictionaries) editing,
    knowledge-base (including transfer-dictionary and
        terminology-bank) construction and maintenance,
    translating processes including post edit,
    text processing.
d. Main computer system is DEC 20 and planned to VAX 11/780.
e. Basic softwares are MacLisp, EMACS modified to treat KANJI (Chainese characters), and Edinburgh's Prolog.
f. The project team has eight members.
g. Prototype system is now on implementing.
h. LDOCE, Japanese and English-Japanese dictionaries are on computer.

2. Goals

a. High quality translation from Japanese into English and vice versa.
b. Corpus is scientific articles (ex. Scientific American for English into Japanese, and "Saiensu" for Japanese into English. "Saiensu" is Japanese and equivalent word to "Science" in English).
c. Totally intelligent systems with good man-machine interface.
d. Preparing the grobal linguistic data and knowledge-base.
e. Much interest is on the role of knowledge in understanding.
f. Knowledge acquisition capability through experience and interaction with human.

3. Approaches

a. Transfer approach.
b. Case structure (including modality) analysis.
c. Discource analysis.
d. Structural transformation.
e. Intermediate structure as an understood state. It contains both of meanings and surface structures for high quality but faithful translation.
f. Intermediate Structure consists of three parts:
   Topics,
   Discource,
   Parsed Structures of Sentences.
g. Frame-like uniform representation of Case frame, Intermediate Structure, Dictionaries, and Knowledge-Base.
h. Man-Machine interactions via high resolution color multi-window display.

4. Current state of research

a. Linguistic analysis of English articles from "Scientific American" written in English by English native speakers.
b. Linguistic analysis of Japanese articles from "Saiensu" written in Japanese by Japanese native speakers.
c. Linguistic analysis of original and translated articles which have same contents. Original articles are written by native speakers.
d. Editing system with high resolution display is under implementation.
e. An experimental translation system is under implementation and planned to operate at next spring.