NEC'S MACHINE TRANSLATION SYSTEM "PIVOT"

Toshio Harada
Manager
EDP Systems Engineering Div.
Application Systems Dept.
NEC Corporation

INTRODUCTION

NEC's PIVOT system is a bi-directional (Japanese-to-English and English-to-Japanese) machine translation system that makes use of artificial intelligence technology and an extensive knowledge base containing expert knowledge related to the subjects and/or fields to be translated. The PIVOT system employs the pivot language method, a revolutionary machine translation method, which, unlike the more widespread transfer method, uses intermediate linguistic representations based on concept structures that are not language-dependent. The pivot language method of machine translation makes it possible for the PIVOT system to produce high quality translations from Japanese to English and vice versa, and will facilitate the expansion of PIVOT to a multilingual machine translation system in future (See Figure 1). Five special features of the PIVOT system are described below.

High-quality Translations Using AI Technology

To achieve high-quality machine translations, a system must be equipped with expert knowledge on all the subject areas and fields it will be tasked to produce translations in. The PIVOT system has such expert knowledge models stored in an extensive knowledge base, which it can refer to as needed when analyzing a source language text and/or generating a target language text.

Bi-directional Translation Capabilities Using the Pivot Language Method

The PIVOT system makes use of a translation process that analyzes the source language text, then, based on those results, constructs a language-independent conceptual structure that serves as an intermediate linguistic representation upon which to generate the target language text. This approach promises to make it quite easy to expand the PIVOT system into a mutllingual machine translation system in future.

Menu-driven Operation for Ease of Use

All the human operations related to the machine translation process, from pre-editing (source language text input
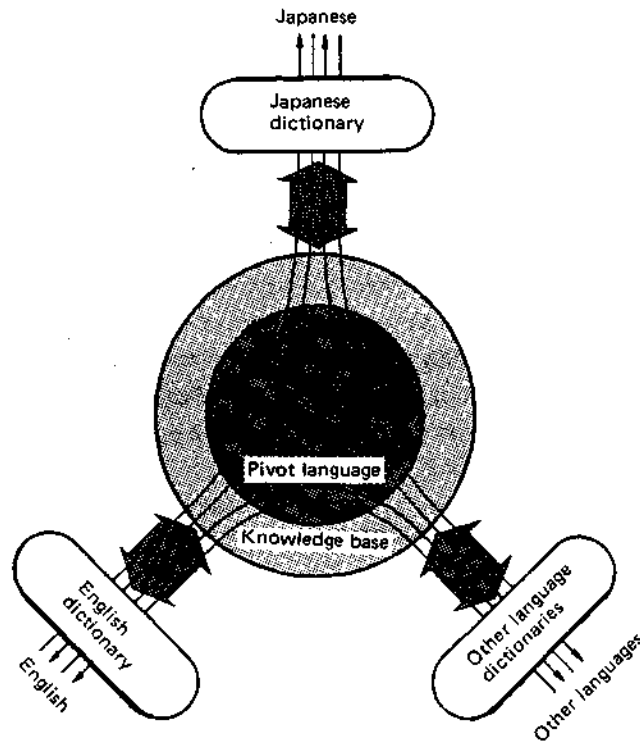
Figure 1. Schematic Diagram of PIVOT System Showing
Its Knowledge Base and Pivot Language

and correction) and translation imple-
mentation, to post-editing (correction
and printing out of target language text)
and the preparation and updating of dic-
tionaries, are made easy by means of
menu-driven screens that guide the
operator through each stage of the
process.

Interconnectability with Integrated Office Systems

The PIVOT system is equipped with
interface capabilities that enable it to
be interconnected to integrated office
systems, such as multi-media document

processing systems and/or electronic
mail/filing systems.

Batch and Interactive Processing Modes

PIVOT can be operated in either the
batch or interactive processing modes.
The former is used to translate large
volumes of material automatically,
whereas the latter mode is used to trans-
late smaller texts using the interactive
approach by which the operator works
together with the system to complete the
translation. In the batch processing
mode, source texts can be input directly

from magnetic media, via optical character readers (OCRs) or from the disk drive devices of word processors on which the document to be translated has been prepared. The system then translates the text on its own. The interactive processing mode, however, allows the operator to input the source language text from a workstation and then work with the machine in the performance of the actual translation. The interactive mode also permits voice input of source language texts.

PIVOT OPERATION PROCE-DURES

The standard operating procedure for the PIVOT system consists of seven steps: source language text preparation, preliminary operations, pre-editing, translation, post-editing, target sentence extraction and target language text print-out. Each of these procedures is explained briefly below.

Source Language Text Preparation

The first step in the machine translation operation procedure for PIVOT is the preparation of the source language text, using the word processing functions built into the system. These word processing functions are provided via NEC's "LANWORD," a document preparation program for use on workstations.

Preliminary Operations

Preliminary operations include specifying the file or files that contain the source language text, determining those words in the source text that have not yet been defined and registering them in the dictionary.

Pre-editing

Pre-editing consists of correction and editing operations performed on the source language text to enhance its translatability.

Translation

As explained above, the PIVOT system offers two modes of translation, batch and interactive. When the interactive mode is used, both the source language text and target language text are displayed on the operator's screen simultaneously so that he/she can work with the system to produce a polished translation the first time around.

Post-editing

Post-editing is performed on the target language text on an interactive basis using built-in word processing functions.

Target Sentence Extraction

Target sentence extraction is the process used to extract the completed translation of the target-language sentence from the system and output it to the word processor file, where it is formatted in preparation for printout.

Target Language Text Print-out

The last step in the PIVOT system's operation procedure is the printing out of the target language text on the host computer or workstation printer.

All of the above cited steps are menu-driven, and require no special knowledge on the part of the operator. And a variety of utility programs, related primarily to the dictionaries, are provided to support PIVOT.

PIVOT TRANSLATION PROCESS

PIVOT employs the revolutionary pivot language method of machine translation, and is equipped with an extensive knowledge base, which the system references during the translation process. Figure 2 shows the translation process utilized by PIVOT. As you can see, the PIVOT system is divided into two major elements: source language text analysis and target language text generation. Each of these elements comprises three modules, all six of which are briefly explained below.
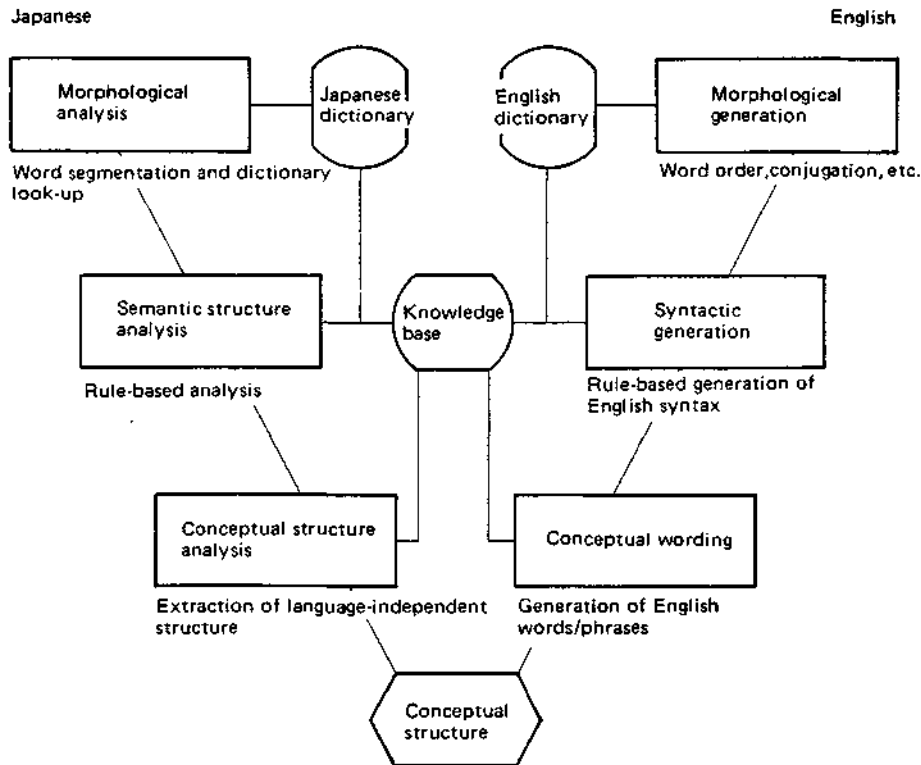
Japanese   English

Figure 2. PIVOT Translation Process

Morphological Analysis Module

The morphological analysis module is the first of the source language text analysis modules, and is tasked with the jobs of segmenting input sentences into words, designating the parts of speech of each word and processing word dependency relations.

Semantic Structure Analysis Module

The semantic structure analysis module does just that, it analyzes the semantic structure of the source language text using an Augmented Dependent Grammar. When analysis is complete, the module constructs a semantic structure tree containing all the words used in the sentence.

Conceptual Structure Analysis Module

The conceptual structure analysis module extracts the conceptual structure from the semantic structure of the source language text. This conceptual structure might also be called the universal language structure.

Conceptual Wording Module

The conceptual wording module is the first module in the target language text generation element. This module converts the conceptual structure of the source language text into the conceptual wording appropriate for the target language text.

Syntactic Generation Module

The syntactic generation module applies sentence generation rules to generate the syntax for the target language text based on the semantic structure tree.

Morphological Generation Module

The morphological generation module is the third and final module in the target language text generation element, and is used to generate the morphology of the target text and to output the final target language text as a character string in accordance with word-ordering information.

The conceptual structure approach to machine translation used in the PIVOT system is expected to make the transition up to a sentence-based (as opposed to the current word-based) translation process relatively easy. It will also make it easy to upgrade the present bi-directional system to a multilingual one in future.

PIVOT SYSTEM CONFIGURATION

The PIVOT system comprises a mainframe host computer and a workstation (See Figure 3), which are under the control of one of two NEC proprietary general-purpose operating systems, the ACOS-6/MVX or the ACOS-4/MVP XE. Host computer-workstation connection

and interaction are performed by a program called Micro-Mainframe Link Software. The workstation is used to prepare and pre-edit source language texts, to perform interactive translation and editing and to post-edit target language texts. The host computer provides the computing power necessary to perform the complicated task of translation processing. Table 1 lists and explains the various translation software programs implemented on the host computer and workstation, respectively.

BENEFITS OF USING THE PIVOT SYSTEM

PIVOT can be used to perform translations on a wide range of different topics, everything from technical documents and manuals to business proposals, cost estimates, patents and contracts. And all these different types of translations will be performed rapidly and efficiently. The PIVOT system reduces human translator idiosyncracies and provides the user with translations that are uniform in quality. And interconnecting PIVOT with office, printing or other pertinent systems makes it possible to create a number of integrated systems with built in translation capabilities.

Table 1. PIVOT Software Programs

|	Program	Explanation
Host programs	TRANSLATOR-JE	Performs Japanese-to-English translation (Processes undefined words and performs batch and interactive translation)
	TRANSLATOR-EJ	Performs English-to-Japanese translation (Processes undefined words and performs batch and interactive translation)
	SDIC-J	Dictionary of basic Japanese terminology
	SDIC-E	Dictionary of basic English terminology
	DICEDIT	Updates dictionaries (Terminology registration, correction and deletion)
Workstation programs	TRANSLATOR-JE	Performs Japanese-to-English translation (Source language text input, interactive translation, target language text editing and final print-out)
	TRANSLATOR-EJ	Performs English-to-Japanese translation (Source language text input, interactive translation, target language text editing and final print-out)
	LANWORD	Word processing software
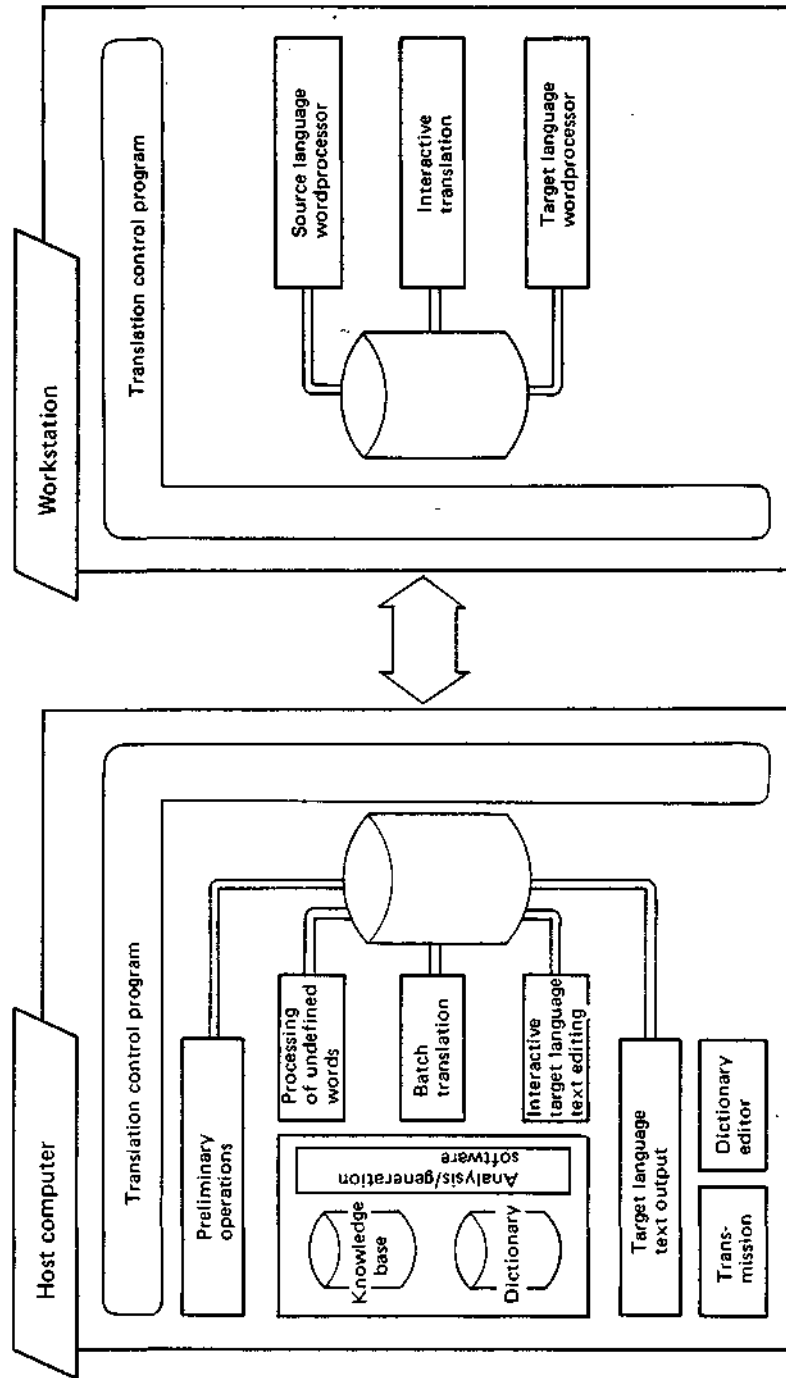
Figure 3. PIVOT System Configuration

国際社会での日本の地位が向上し、日本語の文章を大量に翻訳する必要が生まれてきた。機械翻訳システムの開発は、社会の関心と期待に支えられ、急ピッチに進んでいる。日本電気は日英機械翻訳システムを開発した。このシステムにおける、方式上の特徴は構文に関する情報と意味に関する情報を分離した点にある。構文情報は文法規則の形で持ち、意味情報は知識ベースに格納する。また、言語に依存しない中間表現を介して翻訳を行うため、日本語文の解析と英語の生成の二つの過程を分離している。この翻訳方式は、ピボット方式と呼ばれ、多言語間翻訳を指向している方式である。さらに、知識ベースを利用することにより高品質翻訳を可能にしている。こうした工夫によって、システムの能力は向上している。

The position of Japan in the international society improved, and necessity of translating the text in Japanese in large quantity appeared. Interest and expectation of society supports and development of the machine translation system proceeds rapidly.

NEC developed the Japanese to English machine translation system. In this system, the methodological characteristic is located in the point where information on syntax and information on meaning have separated. Syntactic information is held as the form of the grammatical rule, and conceptual information is stored in the knowledge base.

Two processes of analyzing the Japanese text and generating English are separated in order to translate through the internal representation which does not depend on the language.

This translation method called as the pivot method is the method aiming at multi-lingual translation.

Moreover, high quality translation is made possible by using the knowledge base.

Such contrivance improves performance of the system.

Figure 4. Sample Translation Produced by PIVOT System