PRACTICAL EXPERIENCE OF MACHINE TRANSLATION
V. Lawson (ed.)
North-Holland Publishing Company / ©ASLIB, 1982

PRACTICAL EXPERIENCE OF MACHINE TRANSLATION

Stanley P. Sereda

General Motors of Canada Limited (CANISCA)
Oshawa, Ontario
CANADA

Post-editing is one of the most significant factors in the operation of a computer translation system. The economic validity of computer translation stands or falls on the efficiency and success of the post-editing process. The factors affecting the post-editing functions include the linguistic performance of the system, the quality of the source text, availability of terminology, capabilities of the personnel and mechanical aspect of the translation process.

Good morning ladies and gentlemen. I would like to express my sincere appreciation on behalf of General Motors of Canada and myself for the honor of participating in what I believe will be a most informative and educational conference. The exchange of knowledge is a vital necessity in order to secure progress in our specialized field of machine translation.

General Motors of Canada became actively involved in language translation using a computer in 1976 with the installation of a system called Systran. It was installed with the purpose of translating large volumes of technical literature such as the vehicle service manuals from English to French. Since then we have expanded the application to include railroad locomotives and highway transit coaches.

We are currently staffed with three English-French bilingual translators, one English-French-Spanish trilingual translator and one English-French bilingual typist. In addition one data processing programmer is responsible for the ongoing maintenance of the system.

Since the system has been installed, we have translated the following documents:

700 pages of the 1976 Chevrolet truck shop manual for the Canadian Department of National Defence

350 pages of a locomotive blower-type diesel engine manual

350 pages of a locomotive turbo-charged diesel engine manual

448 pages of a locomotive service manual

118 pages of a locomotive operator's manual

450 pages of a transit coach manual

100 pages of an Isuzu owner's manual.

And on a continuing basis, we are translating the following items:

the Product Service bulletins (averaging 72 pages per month)

technicians training guides (averaging 50 pages per guide).

Both the Electro-Motive diesel engine shop manuals and the Department of National Defence shop manual have been produced in a two-column side by side bilingual format.

In discussing language translations using a computer, one very important point must be noted. No computer translation system, as you are aware, is perfect. Due to the intricacies in language rules the system cannot be expected to produce error free translation. Therefore, it must be understood that a computer translation system is in fact a computer assisted translation system where the human translator and not the computer plays the key role.

The function of the translator in the computer environment differs from the manual environment in that the translator becomes a post-editor devoting his time and attention to refining the translation rather than spending a great deal of tedious time on manual translation of common words. Ideally, the translator's function should be to proof read the computer translated text and make few necessary refinements. This will happen only through vast improvements in the computer translation technology.

With today's technology, the post-editing function is the most time consuming and costly segment of the computer translation process. In order to make the system economically viable, the post-editing function must be efficient to realize sufficient savings over the manual process to cover the extra cost of computer resource usage.

On any given day, an experienced translator working on technical material, can manually produce final copy at a rate varying between 800 and 1,500 words per day depending on the difficulty of the text. A computer will process the same work in less than two minutes. Currently, our translators are able to post-edit at a rate of 3 to 4 times faster than manual translation. We believe this ratio can be increased further with linguistic enhancements.

The factors that determine the effectiveness of computer translation and specifically the post-editing function can be classified as follows:

- the linguistic performance of the system
- the source language text to be translated
- the availability of terminology
- the translators who carry out post-editing

and

- the mechanical aspect of the system.

It is obvious that the linguistic performance of the computer translation system is a vital factor in determining the efficiency of post-editing. If the analysis and synthesis systems function incorrectly, then the target language text will be difficult to edit. Experience has shown that a simple word-for-word translation is impractical to post-edit; that is, the cost of machine translation is greater than the cost of manual translation. The translation system must carry out a certain level of "intelligent" analysis of the source language and selective synthesis of the target language.

It should be noted that different errors in translation have a different effect on the post-editing function. Minor errors involving articles and verb/adverb re-arrangement can be resolved quickly and easily by the translator. On the other hand, certain kinds of structural errors can be extremely difficult to correct, perhaps requiring the complete rewrite of the affected sentence.

The linguistic performance of the system is the main factor that affects post-editing. Unfortunately, it is the only area in the translation process where the user of the system has very little control in implementing necessary improvements. User can only identify the problem areas that require correction and then wait for the errors in the system to be corrected.

Inconsistent linguistic analysis, unless corrected on a timely basis, can cause a great deal of frustration on the part of the translator. Of course, we realize that it is impossible to correct all linguistic problems. However, we hope that further enhancements to the system will minimize errors thus reducing the post-editing requirement.

The source text affects the degree of post-editing required in the translation. Text containing grammatical errors or that bends the rule of the source language will produce unpredictable translation. For instance, some of the problems encountered in technical publications are incomplete sentences, ambiguous text due to lack of articles and punctuation and the use of abbreviations. Well written, well punctuated and unambiguous source text results in the translation that will require minimal post-editing. The use of consistent terminology and sentence structure will also lessen the need for post-editing.

To achieve acceptable level of translation, it may be necessary to evaluate and pre-edit the source text in a way that recognizes the limitations of machine translation. The need for pre-edit can be further reduced through controlled writing by applying certain text preparation guidelines on terminology usage and sentence structure.

It is possible to lessen the overall translation requirements, especially with technical manuals, by substituting illustrations in place of the written text while the source materials are being prepared.

It is apparent that even the most advanced computer translation system will be useless without the availability of sufficient terminology in the source and target languages. Our entire vocabulary is contained in two dictionaries, stem dictionary which contains single words and a dictionary containing multi-word expressions. Currently these dictionaries contain 52,500 stem words and 78,000 expressions which are constantly being updated. Most of the words and expressions in the dictionary are technical terms. The vocabulary is being expanded to include words and expressions pertaining to other subject fields of our business. We are currently working to expand the English-French vocabulary to include Spanish. The English-Spanish dictionary currently contains 14,200 stem words and 5,350 expressions.

During the early stages of our activity we undertook an extensive dictionary coding procedure. We would translate a document, then code all unfound words and expressions for updating the dictionaries. As the vocabulary increases, the volume of the dictionary update decreases. However, the time and effort required for terminology research do not decrease proportionally with the volume. Finding a correct equivalent for technical terms is not a simple dictionary look-up operation. It is a difficult intellectual process involving knowledge of the technical field and the practices of the "source" and "target" languages.

Clearly, improvements to the dictionary cost time and money, and should be offset by improvements in the post-editing performance. Thus, items which occur very infrequently in real text should not be coded as it is unlikely that the benefit will match the cost of the coding operation. Another point to consider in dictionary coding is that due to the nature of linguistic analysis, a dictionary change made for a specific text can have a negative effect on the translation of other texts.

One factor in the terminology research problem is perhaps peculiar to North America. Because of the cultural and technological dominance of the English language, the English terminology especially in technical fields is pervasive even in the francophone Province of Quebec. This situation in Quebec is expected to change through the program of francization undertaken by the Provincial Government of Quebec.

The process of post-editing involves two main tasks; to identify errors in the translation and to find solutions for the errors. To carry out these tasks, the translator must be completely fluent in the target language as well as the source language. Furthermore, the individual should have a good knowledge of the subject matter involved in the translation to be an effective post-editor. An individual fluent in the target language will be able to recognize linguistic errors with little difficulty. But the translator will have difficulty detecting meaning errors unless he can read and understand the source text. In technical text, the understanding of the subject matter will further enhance his ability to identify factual errors.

Once an error has been identified (whether a linguistic error or a meaning error), a correct form must be found. This process also requires native knowledge of the target language and specific technical skills in the appropriate field. It should be noted however that the translator must be controlled to some extent - particularly in terms of using acceptable standardized terminology and in terms of not wasting excessive amounts of time on purely stylistic changes.

It is possible that this phase involves the translator in consultation with other technical specialists or reference material.

The availability of the computer resources for translation activity offers the opportunity to take advantage of these resources to facilitate the post-editing function.

At General Motors, text processing and word processing systems are used to pre-edit the source text and to post-edit the target text. User friendly procedures have been developed to allow the translators to initiate computer translation by specifying options and parameters via video display terminals. The translators are provided with a document which lists the source and target texts in a side by side format to post-edit the translation.

To simplify terminology look-up, we have developed an English-French on-line dictionary system. This system provides the instantaneous translation of words or expressions contained in our dictionaries on a video display screen, in their basic forms, eliminating the need for hard copy dictionary listings.

We have also developed a terminal entry dictionary coding system to assist translators in updating the dictionaries. This codingsystem eliminates the need for hard copy coding sheets and at the sametime allowsus to control job submissions in a more efficient manner.

We have developed pre-processor and post-processor programs to simplify the production of the target language document. Any data such as photocomposition codes and text processing codes if left in the text may cause erroneous translation. The pre-processor program flags these codes as "do not translate" thus eliminating these codes from the translation process. The post-processor reintroduces these codes back into the target language text thus eliminating the need for re-keyboarding them.

As you can gather, we have endeavoured to mechanize and facilitate the post-editing process to maximize the translation productivity. However, our experience shows the need for continuing improvement in the linguistic performance of the

translation system. These improvements are of great importance to us because as General Motors' marketing of vehicles and plant facilities expand world-wide, we anticipate our translation requirements will expand dramatically.

We believe that the computer assisted translation system will play an integral part in meeting this translation demand.

ELECTRO-MOTIVE

ENGINE MAINTENANCE MANUAL
MANUEL D'ENTRETIEN DU MOTEUR

Section 5

CYLINDER POWER ASSEMBLY

DESCRIPTION

Sections 2, 3, and 4 contain information on the cylinder head, piston and connecting rod, and the cylinder liner respectively. Procedures are provided in these sections for disassembly and assembly of the power assembly components beyond what is done during removal from and installation into the engine. Also, the information concerning cleaning, inspection, and the qualification of components is detailed in these sections.

The following procedures are for the removal and installation of a cylinder power assembly, component by component, and the removal and installation of the power assembly as a unit.

COMPONENT BY COMPONENT REMOVAL

1. After draining the cooling system, remove the top deck cover over the affected cylinder. It is advisable to remove the front latches first, then the rear latches.
2. Remove the air box and oil pan handhole covers for the cylinder being removed and the opposing cylinder on the other side of the engine.
3. Remove the piston cooling oil pipe.
4. Remove the bolts holding the water inlet tube to the cylinder liner and remove the saddle strap nuts holding the tube to the water manifold.
5. Remove the gasket from the water manifold.
6. Open all cylinder test valves using the test valve wrench. This will facilitate manual barring of the engine.
7. When removing a fork rod assembly, bar the engine over until the piston is 120° after top dead center. This will allow removal of the basket halves and the connecting rod bearing shells at one crankpin position.
8. Loosen the cylinder test valve packing nut and remove the cylinder test valve and seal. The entire test valve assembly must be removed before removal of the head, or damage to the head and/or the test valve will occur.

GROUPE-MOTEUR DES CYLINDRES

DESCRIPTION

Les sections 2, 3, et 4 contiennent des renseignements sur les culasses, les pistons, les bielles, et les chemises de cylindre respectivement. Dans ces sections, on donne les méthodes de démontage et de montage des éléments du groupe-moteur ce qui est au-delà de ce qu'on fait lors de leur pose et dépose dans le moteur. Aussi, des renseignements concernant le nettoyage, l'inspection, et l'admissibilité des éléments sont détaillés dans ces sections.

Les méthodes suivantes couvrent la dépose et la repose du groupe-moteur, pièce par pièce, ainsi que la dépose et la repose du groupe-moteur comme unité.

DÉPOSE PIÈCE PAR PIÈCE

1. Après la vidange du système de refroidissement, enlever le couvercle du pont supérieur du cylindre affecté. Il est recommandé de libérer les loquets avant d'abord, puis les loquets arrière.
2. Enlever les couvercles de visite de la boîte d'air et du carter d'huile du cylindre affecté et du cylindre opposé situé de l'autre côté du moteur.
3. Enlever le tuyau d'huile de refroidissement du piston.
4. Enlever les boulons retenant le tube d'arrivée d'eau à la chemise du cylindre et enlever les écrous du collier d'attache retenant le tube à la chemise d'eau.
5. Enlever le joint d'étanchéité de la chemise d'eau.
6. Ouvrir les soupapes d'essai de tous les cylindres en utilisant la clé pour soupapes d'essai. Ceci facilitera la rotation manuelle du moteur.
7. Lors du démontage d'une bielle à tête fourchue, faire tourner le moteur jusqu'à ce que le piston soit à 120° après le point mort haut. Ceci permettra la dépose des demi-godets et des coussinets de bielle de certains manetons.
8. Desserrer l'écrou du presse-étoupe de la soupape d'essai du cylindre; enlever le joint d'étanchéité et la soupape d'essai du cylindre. La soupape d'essai doit être enlevée en totalité avant la dépose de la culasse, sinon il y aura dommage à la culasse et/ou à la soupape d'essai.

(FRENCH)

SERVICE DEPARTMENT

ENGINE MAINTENANCE MANUAL

MANUAL DE MANTENIMIENTO DEL MOTOR

Section
Sección
5

CYLINDER POWER ASSEMBLY

DESCRIPTION

Sections 2, 3, and 4 contain information on the cylinder head, piston and connecting rod, and the cylinder liner respectively. Procedures are provided in these sections for disassembly and assembly of the power assembly components beyond what is done during removal from and installation into the engine. Also, the information concerning cleaning, inspection, and the qualification of components is detailed in these sections.

The following procedures are for the removal and installation of a cylinder power assembly, component by component, and the removal and installation of the power assembly as a unit.

SPECIAL TOOLS

The special tools required are listed in the Service Data at the end of the section.

COMPONENT BY COMPONENT REMOVAL

1. After draining the cooling system, remove the top deck cover over the cylinder affected. It is advisable to remove the front latches first, then the rear latches.
2. Remove the air box and oil pan handhole covers for the cylinder being removed and the opposing cylinder on the other side of the engine.
3. Remove the piston cooling oil pipe.
4. Remove the bolts holding the water inlet tube to the cylinder liner and remove the saddle strap nuts holding the tube to the water manifold.
5. Remove the gasket from the water manifold.
6. Open all cylinder test valves using the test valve wrench. This will facilitate manual barring of the engine.
7. When removing a fork rod assembly, bar the engine over until the piston is 120° after top dead center in the cylinder being removed. This will

CONJUNTO DE POTENCIA DEL CILINDRO

DESCRIPCIÓN

Las secciones 2, 3 y 4 contienen información sobre la culata del cilindro, los pistónes, la biela y la camisa del cilindro respectivamente. Los procedimientos ofrecidos en estas secciones son para la remoción y la instalación de los componentes del conjunto de potencia más allá de lo que se realiza durante la remoción y la instalación en el motor. Además, la información referente a la limpieza, la inspección y la calificación de los componentes se detalla en estas secciones.

Los procedimientos siguientes son para la remoción y la instalación del conjunto de potencia del cilindro, componente por componente, y la remoción y la instalación del conjunto de potencia como una unidad.

HERRAMIENTAS ESPECIALES

En los Datos de Servicio en el fin de la sección se ofrece una lista de las herramientas especiales requeridas.

REMOCIÓN COMPONENTE POR COMPONENTE

1. Después de drenar el sistema de enfriamiento, retirar la tapa de la plataforma superior sobre el cilindro en cuestión. Se aconseja retirar primero los pestillos delanteros, y luego los pestillos de la parte posterior.
2. Retirar las tapas de los orificios de limpieza de la caja de aire y del colector de aceite del cilindro que se está retirando y del cilindro opuesto en el otro lado del motor.
3. Retirar el tubo de aceite de enfriamiento del pistón.
4. Retirar los pernos que sujetan el tubo de entrada de agua a la camisa del cilindro y retirar las tuercas del collar de fijación que sujetan el tubo al múltiple de agua.
5. Retirar la junta del múltiple de agua.
6. Abrir todas las válvulas de pruebas del cilindro usando la llave de la válvula de pruebas. Ésto facilitará la sujeción manual del motor.
7. Al retirar el conjunto de la biela de horquilla, sujetar el motor hacia arriba hasta que el pistón esté a 120° después del punto muerto superior en

(SPANISH)