Title : Machine Translation in Post-Contemporary Era

Author name(s) : Grace Hui Chin Lin

Adjunct Professor, National Changhua University of Education

Publication date: (or date of document completion)
2010, Dec. 10

Conference information for conference papers
(name, date and location of the conference) :

Name : 2010 International Conference on TESOL and Translation, Dept. of English Language, Da-Yeh University

Date: Dec. 10, 2010,

Location: Changhua, Taiwan

Machine Translation in Post-Contemporary Era

Hui-chin Lin

National Changhua University of Education

Abstract

This article focusing on translating techniques via personal computer or laptop reports updated artificial intelligence progresses before 2010. Based on interpretations and information for field of MT by Yorick Wilks' book, *Machine Translation, Its scope and limits,* this paper displays understandable theoretical frameworks from views of a translating field worker and linguistics. Moreover, from author's practical application experiences working as a translator, this research in addition introduces two current and popular soft wares and translating systems created by SYSTRAN and Google. The basic functions and characteristics created by computering assistant translation are thoroughly examined and introduced. More importantly, the advantages and weaknesses in computational assisting procedures are investigated and revealed. The purpose of this study mainly focuses on providing an overview for what functions and what doesn't perform well in computational linguistics, comparing to human translation. After reading this commentary-oriented article, the reader will obtain basic concepts for definitions and explanations for diverse terminologies in MT field.

Key words: Machine Translation, Artificial Intelligence, Computational Linguistics

Introduction

Machine Translation has been a significant issue in our post-modernized world where almost everyone is able to use a computer and surf on the internet. In fact, the field of Machine Translation (MT) has been developed for about 40 years, according to Yorick Wilks (2009), a professor focusing on his studies for artificial intelligence and the computer processing of language and Semantic Web. Therefore, there have been diverse softwares of MT can be found either in 3C (computer, communications, and consumer electronics) stores or on-line web stores. For example SYSTRAN is a type of popular software sold in the market and Google translator belongs to a free on-line processor.

Machine Translation belongs to a type of simultaneous translation. By imitating this system in human translator's brain, electronic translation formats were designed. Form views of a translating field worker and computational linguists, there have been several process models of simultaneous interpretation drawn by scholars (e.g., Gerver, 1976; Moser, 1978). The concurrent interpretation by machine's function mostly had been demonstrated by a series of processes how foreign language resource input are transformed into translated language output, including steps of decode and store or decide related syntactic/semantic context information. Among several models declaimed, it can be discovered that the electronic cognizant functions of MT primarily are memorization, simultaneous interpretations as well as instantaneous corresponding translating, stronger lexical data driven and weaker grammatical rule driven features. Hence, human being double checking for editing sentence structures, logic meanings and correcting fossilized data without following the grammatical rules are definitely needed.

Literature Review

In Yorick Wilks (2009) book *Machine Translation, Its Scope and Limits,* a long-term surviving system, SYSTRAN of MT has been mentioned. In fact, SYSTRAN was founded half decade ago by Dr. Peter Toma in 1968, which is one of the most senior MT companies in the world. So far, it has effectively and practically achieved extensive translation works for the United States Department of Defense and the European Commission. In order to know more about this updated MT software when keying title of SYSTRAN and searching for its advertisement webpage, its price and explanations for specific functions can be discovered in a second. Since there is a huge market selling translating software in our globalised world where communications are frequently, this prestigious and well programmed MT designing industry is in fact a large company and currently already has about 55 employees. Most of the employed are computational experts and computational linguists in professional fields of computer science and language studies. According to Wilks (2009), " …about 60% of sentences correctly translated from SYSTRAN…output that full MT system continue to produce and which is used for rough draft translations" (p. 6) The rest 40% part of MT texts currently still need human eyes to clarify them in order to obtain a clear and understandable translated texts. Nevertheless, for professional translators, having a draft roughly interpreted before further editing and revising is still very valuable and needed. Naturally, translation field workers would agree to purchase it and also anticipate that more humanized MT system can be developed sooner by computational linguistics developed by upgraded the software.

Besides expensive softwares that need to be purchased and exploited by significant associations such as United Nation or World Trade Organixation, some general translating softwares can be easily found on the World Wide Webs. For example, when the author of this article worked as a post-doc researcher as well as translating field workers for professors at Department of Electronics and Engineering

of National ChunHsing University, Taichung and Department of Library Information Science of National Taiwan University, Taipei, Google translator for example, can be easily applied for translating academic theses, including professor's Chinese publications and graduate students' home works. From perspectives of a field worker or and a translation methodology researcher, the role of MT as the author see it cannot would be just a theoretical updated issue under the situation that we just learn the theory of Translation subject instead of applying the machine and software for creating economic worth. Indeed it if it is applied practically and efficiently in field of commercial markets or academic theses translation, it becomes a very much benefit-productive and really precious instrument of making a life.

Advantages and Weaknesses in Computational Assisting Procedures

Based on translating needs of trade company's commercial letters, academic texts or scholar's studies, the functions of MT can be implemented in an outfitted and beneficial approach, according to the author's translating experiences working as a translation post-doc researcher at 2 schools. Although ambiguous translated texts can be not avoided after all, countless scholars still intend to declare its advantages. "It is now widely accepted that global communications just be accessible and transferable, in a timely manner, in as many languages as feasible." (Açıkgöz & Sert, 2006, p. 1) Nevertheless, there is still a huge space for human being to correct and humanize for the translated output. In truth, necessary translation strategies and training for how to organize accurate sequences of translated lexical items and data output associated with pragmatics, semantic and sociolinguistics are needed by a professional translator. Also, this might be a long-term defect of MT that will be never resolved.

Açıkgöz and Sert (2006) two Turkish linguistics have investigated the literatures discussing the resolutions for MT problem and provided sufficient reports in their study. They note that most authors of the literature on MT systems have

focused on a particular problem and stated the limitations of their studies. First of all, Volk (1998) in his study titled *The automatic translation of idioms-translation memory system Vs. machine translation* has motioned the impracticability of translating idioms due to syntactic setbacks, but proposes an untested and questionable approach, whereas the truth that language cannot be separated from its segments either culturally or linguistically remains disregarded.

> "Translating idioms is one of the most difficult tasks for human translators and translation machines alike. The main problems consist in recognizing an idiom and in distinguishing idiomatic from non-idiomatic usage. Recognition is difficult since many idioms can be modified and others can be discontinuously spread over a clause. But with the help of systematic idiom collections and special rules the recognition of an idiom candidate is always possible. The distinction between idiomatic and non-idiomatic usage is more problematic. Sometimes this can be done by means of special words that are only used in an idiom." (p. 167)

What's more, with a parallel attitude, Arnold et al. (1995) have also interpreted for above issue related to MT training, which suggests a comprehensive assessment of the issues behind MT and frequent misconceptions.

Conclusions

To conclude, for clarifying field of Machine Translation, two groups of scholars should be divided and defined. They are Translation scholars discussing the usages and functions developed by MT industries and Machine users as well as the practical translators implementing the developed software in field working. On the other hand, there are the other group working as computational linguists. Computational linguistics can be defined through detailing upon the involvement of linguists, computer scientists, experts in artificial intelligence, mathematicians, logicians,

cognitive scientists, cognitive psychologists, psycholinguists, anthropologists and neuroscientists. In other words, the author would like to divided the field workers by in side of industry, as well as the designers and outside of the science park, as well as the users. The translating tasks of the users and translators are more general and common and the designing/upgrading tasks need deeper training and professional knowledge in diverse field of Linguistics, Sociolinguistics, Computer Science,...etc. Summing up, if MT needs to be deeply and professionally interpreted, this tool of translating foreign language works can be defined from views of computer programming, assessories creating, language generating and organizing, inter-cultural differences, standard linguistics rules, and the working processes of utilizing it.

Reference

Açıkgöz, F. & Sert, O. (2006). Interlingual Machine Translation: Prospects and Setbacks. http://accurapid.com/journal/37MT.htm

Arnold et al. (1995). *Machine translation-An introductory guide,* Oxford, UK: NCC Backwell Ltd.

Gerver, D. (1976) Empirical studies of simultaneous interpretation: A review and a model, in Brislin, R. (ed.) 1976: *Translation*, New York: Gardner Press, 165-207.

Moser, B. (1978) Simultaneous Interpretation: A Hypothetical Model and its Practical Application, in Gerver, D. and Sinaiko, W.H. *Language, Interpretation and Communication*. London, Plenum.

Volk, M (1998). *The automatic translation of idioms. Machine translation Vs. translation memory systems.* In: Weber, Nico. Machine translation: theory, applications, and evaluation. An assessment of the state of the art. St. Augustin, 167-192.

Wilks, Y. (2008). *Machine translation: Its scope and limits.* New York: Springer.