Machine Translation for Academic Purposes

Grace Hui-chin Lin

PhD Texas A&M University College Station

Master of Science, University of Southern California

Paul Shih Chieh Chien

Center for General Education, Taipei Medical University

Abstract

Due to the globalization trend and knowledge boost in the second millennium, multi-lingual translation has become a noteworthy issue. For the purposes of learning knowledge in academic fields, Machine Translation (MT) should be noticed not only academically but also practically. MT should be informed to the translating learners because it is a valuable approach to apply by professional translators for diverse professional fields. For learning translating skills and finding a way to learn and teach through bi-lingual/multilingual translating functions in software, machine translation is an ideal approach that translation trainers, translation learners, and professional translators should be familiar with. In fact, theories for machine translation and computer assistance had been highly valued by many scholars. (e.g., Hutchines, 2003; Thriveni, 2002)

Based on MIT's Open Courseware into Chinese that Lee, Lin and Bonk (2007) have introduced, this paper demonstrates how MT can be efficiently applied as a superior way of teaching and learning. This article predicts the translated courses utilizing MT for residents of global village should emerge and be provided soon in industrialized nations and it exhibits an overview about what the current developmental status of MT is, why the MT should be fully applied for academic purpose, such as translating a textbook or teaching and learning a course, and what types of software can be successfully applied. It implies MT should be promoted in Taiwan because its functions of clearly translating the key-words and leading the basic learners to a certain professional field can be proved in MIT.

Keywords: Machine Translation, Computational Linguistics, Bi-lingual/Multilingual Translating, Open Courseware

Introduction

Machine translation (MT) is a modern method of translation through computer assistance, which is a sub-field of computational linguistics. This academic and professional area associated with both Translation and Computer Science explores the functions of computer software that can translate texts in foreign languages to be readable and comprehensible. MT performs simple replacements with basic translated key words for non-native speakers to understand content in an original foreign language that they need to assess. Using computer translation software, such as Dr. Eye, translation functions in Google toolbar or Yahoo toolbar, complex contents in foreign languages can be comprehended at a glance by people who does not familiar with the languages in the original contents. The MT software professionally and cautiously created, and assisted language experts and linguists to handle linguistic differences in typology, grammatical differences and idioms.

According to Vauquois (1998), MT is aimed at enabling a computer to transfer natural language utterances, or to process a natural language in terms of lexical, syntactic and semantic dimensions. For both of texts or speeches from one language in the original texts to anther language for the receivers, and even for both of explicit and implicit meanings, the MT productions can be found working in an effective way. Vauquois (1998) implies that MT should be an applicable technique of doing translation.

However, some other scholars of translation field hold a different way looking at this issue. Thriveni (2002) has ever emphasized that "one language can not express the meaning of another;…different languages predispose their speakers to think differently…." Thriveni (2002) surely holds a negative attitude to see the translating functions of processors and soft-wares of computers. He insists that cultural

interpretations and recognitions through natural translation by a translator should be a more precise way of doing translation, since the literature and culture senses in the text and the speeches can not be uncomplicatedly revealed by a machine.

Literature Review

The concepts of idea of machine translation can be found in the early 17th century. In 1629, René Descartes proposed the idea of universal language, with equal ideas in different tongues sharing one symbol. In the 1950s, The Georgetown experiment (1954) developed fully-automatic translations of more than sixty Russian sentences into English through computer. The empirical evidences generated by Georgetown experiment of machine translation was a surprising achievement so that this field of MT is established by later scholars and translators who hold a positive attitude toward machine-translation research. After a long term evolvement, translation through computer assistance becomes a mature field so far, and it is a current issue that is still debated by many scholars in the field of translation (e.g., Acikgoz, 2005; Bahar, 2001; Celik, 2003; Furstenberg, 2001; Rodrigo, 2001; Topcu, 2004).

Melby (1996) firmly notes that for some texts, predominantly highly technical texts treating a very specific topic in a rather dry and monotonous style, computers sometimes do quite well. He implies that future computers might be able to make judgments, decisions and choices consistent with non-linguistic knowledge that people frequently refer to in their daily lives. That is to say, when the contents in original tests are not complex and are highly repetitive texts that can be frequently found in the texts of human being's language of daily lives, the computer sometimes can make accurate judgments and help us to approach the needed meaning of original texts in a foreign language.

Acikgoz and Sert (2006) at Hacettepe University in Ankara of Turkey support above ideas and mention in their conference paper titled Interlingual machine translation: prospects and setbacks, indicating that:

> given that any field in which human beings are actively involved requires the knowledge of another field, MT, having a history almost as old as the modern digital computer, emerged as an attempt to overcome the intricacy of "being informed" in a group of offers to sustain communication. (p. 1)

Acikgoz and Sert's (2006) study obviously leads us to consider if one of the reasons why the world is globalized should be the development and result of development of MT. Due to global residents' assessable knowledge from the web-pages of all the nations' websites, universal residents are becoming intelligent in a way that they can integrate other races' thoughts and philosophies into their own cultural and traditional ideologies.

In fact, there are more and more learning approaches that can be found from Open Courseware in highly industrialized nations, such as Open Course-wares in Massachusetts Institute of Technology, Johns Hopkins Bloomberg School of Public Health, Utah State University, and Tufts School of Dental Medicine in the United States of America, and Paris Tech in France. Most of the online courses translated for the foreigners were mostly ever provided in the formal system of education at a certain university. They later became on-line courses after either naturally or through machine being translated into different languages for learners of academic subjects around the world to learn at home without going abroad.

One of the methods how these sufficient online courses are created, translated, and adopted voluntarily by many translators and teachers should resort to contributions of computer assistance. It is suggested that some translations of the

course contents including course descriptions, assignments, handouts and examinations, were actually firstly translated through computing assistance. Then, they are upgraded, edited, revised and given more details by human translators and experts of the professional and academic fields.

For example, the authors of this conceptual and empirical study tried to adopt and translate an online course in Open Open courseware of MIT, 24.03 Relativism, Reason, and Reality, Spring 2005, through bilingual translating with both of machine and natural translations. From Appendix A of this paper produced by translation tools of Google toolbar, it is clear that the texts assisted through computer are readable and understandable. The translator of this study applied the MT tool of computer and translated the text with one minute, and then it was read by human eyes, edited, and given details for vague words in the translations using another ten minutes. For one page of original text, it took only eleven minutes.

Therefore, the translation version of MIT's course is strongly supported by a scholar Lucifer Chu and also by teachers in schools. According to Lee, Lin and Bonk (2007) "An all-volunteer organization called the Open Opencourseware Prototype System (OOPS), headquartered in Taiwan, was initially designed to translate open source materials from MIT OpenCourse Ware (OCW) site into Chinese." (p. 1) This is absolutely a piece of good news for learners in different academic fields because this international website for learners in different areas provided course syllabi, handouts and examination contents for students in online studies for free. Although it does not provide diplomas and certificates for graduates who have taken the online courses, it makes people around the world who have time and budget acquire approaches to be experts in varied academic fields.

Authors' Example of Machine Translation

Bilingual or Interlingual machine translation is an example of rule-based machine-translation approaches designed in the software that can be easily applied by translators. Through this updated method of translation, English texts can be translated and transformed into at least vague Chinese. Translators usually can edit the ambiguous meanings into comprehensible ones using techniques of repetitions and giving details.

According to the researchers' experiences of translating a course in MIT (see Appendix A), the target language, English can be translated and understood easily through the translation functions in Google toolbar and yahoo tool bar. Based on the experiences of the researchers, the translated words by the tool of MT, eighty percent of the translations can be accurately generated by the computer. The translators as well as the researchers of this study, as well as English teachers at university, only has to spend a little time and energy to read and find the vague parts that needs more repeated details and interpretations. The above ideas may imply that each English major of universities of Taiwan can adopt and translate a course easily when they have time after learning at school.

That is to say, it is really trouble-free to gain the honor as a volunteer of the MIT's course adaptor as well as the academic translator. Why the researchers of this study emphasize this issue of adopting MIT's online course to translate is that the translators can learn a lot of knowledge when translating the course. Moreover, a large number of courses were translated into simplified Chinese. If people in Taiwan insist to read and learn in traditional Chinese, much work of translations into traditional Chinese needs to be done.

Humanized Translations

Dictionary-based translating is a translation method that makes the translated article even more humanized since the meaning of each word can be more accurately located. Besides Google toolbar and Yahoo toolbar that are provided by the internet, dictionary entries, such as Dr. Eye can make the words translated by a dictionary system sound more humanized and normal. Since the high-priced translation softwares are designed by linguistics and translation experts, people might find the translated texts are even more truthful and legible than translated web-pages by toolbars of Google or Yahoo.

In order to make the text translated by machine more intelligible, the translation process through re-encoding by human being's eyes needs to be executed. Re-encoding meanings through understanding the target language and examining the machine translated text are two of the most significant steps in making the translations more humanized. The behaviors of doing translation play an important role in the cognitive operation. For decoding the original ideas in the source text into its completely logical entirety, the translator must explain and analyze all the possible ideas involved in the context. This is actually a step of doing translations which requires sufficient linguistic knowledge of the grammar, semantics, syntax, idioms, in the target/ source language. At the same time, a professional translator should be familiar with the traditions, cultures and linguistics knowledge of the receivers in order to make the translation acceptable.

Most importantly, for making the translated contents more cogent and rational, after re-decoding and making the translations coherent by our eyes, more details and interpretations based on an experts' professional knowledge or the translator's research should be added in order to reproduce the intricate ideas that the receivers

might not be able to straightforwardly comprehend. That is, in order to make the complicated source-contexts explicable, more abundant words should be provided with the translation in Chinese than the original source contents. For example, if there is a page of one-thousand-word article in English, a translation of one thousand and five hundred words to three thousand words of Chinese translations should be provided to the receivers in order to make them easily comprehend the original English content.

Conclusion

Based on one of the authors' experience of translating her own dissertation, she found that a translator needs to fully apply the machine translating functions in order to convey the ideas that a scholar has learned in a target language. The knowledge gain in another nation can be transit to another place if the scholars are willing to apply the MT functions and bring learned updated knowledge to their own lands. If a translator insists to use the traditional way of translation through reading and handwriting without using computer, s/he might not be able to catch the speedy tempo of globalization.

Based on Acikgoz and Sert's (2006) assumption that MT is one of the significant reasons why this world is globalized, the researchers of this study suggest MT must be applied. Applying MT in a humanized way makes people outside of the professional field understand a topic in an in-depth level without delay. In sum, achieving deep understandings for the contents of a professional context through applications of MT as a translator, an expert of the professional field or an English major can do might no more be a utopian concept for anyone in the current globalized world. This means through applying computer assistance, professional knowledge is no more non-assessable. Therefore, for improving ourselves as an updated global citizen,

learning online using the MT functions and translating professional texts through computering soft-wares might be better techniques that we should apply, which would assist us to cast away the limitations in time-consuming and expensive translation.

References

Açıkgöz, F. (2005) *Metaphors and non-native tertiary learners: Ametaphorical study of 'Archetypes'-An absurd short story by Natazsa Goerke*. Unpublished Paper presented at Çukurova University: 3rd International Postgraduate Conference

Açıkgöz, F. & Sert, O. (2006). Interlingual Machine Translation: Prospects and Setbacks. *Translation Journal, 10*(3). Retrieved on Nov. 2nd, 2009 from: http://accurapid.com/journal/37mt.htm

Bahar, I.B. (2001). Linguistic relativity and the translation dilemma: Reading between the lines in Malay literatures in English. *Nottingham Linguistic Circular, 16*, 19-29.

Celik, M. (2003). *Learning Stress and Intonation in English*. Ankara: Gazi

Furstenberg, G. et al. (2001). *Giving a virtual voice to the silent language of culture: The Cultural Project Language Learning & Technology. 5*(1), 55-102

Hutchins, W., J. (2003). Machine translation: half a century of research and use. http://ourworld.compuserve.com/homepages/WJHutchins

Lee, M., Lin, M.-F., & Bonk, C. J. (2007, November). OOPS, turning MIT OpenCourseWare into Chinese: An analysis of a community of practice of global translators. *International Review of Research in Open and Distance Learning, 8*(3). Retrieved on Nov. 2nd, 2009 from: http://www.irrodl.org/index.php/irrodl/article/view/463/980 (HTML) http://www.irrodl.org/index.php/irrodl/article/view/463/982 (PDF) http://www.irrodl.org/index.php/irrodl/article/view/463/966 (audio file)

Melby, A.K. (1996). *The Possibility of Language*. Warner John Benjamins Publishing Company.

Rodrigo, E.Y. (2001). Making MT commonplace in translation training curricula: too many misconceptions, so much potential! In: Maegaard, B. (ed.) *MT Summit VIII*. CD-ROM Proceedings of the 8th International Machine Translation Summit - 'Machine Translation in the Information Age.' Santiago de Compostela (Spain), 18th-22nd September 2001. Geneva: 45-49.

Thriveni, C. (2002). Cultural elements in translation: The Indian perspective. *Translation Journal*. 6/1 Retrieved on Nov. 2nd, 2009 from: http://accurapid.com/journal/19culture.htm

Topçu. N. (2004) 'Les locutions idiomatiques imagées dans la presse turque et leur utilisation dan une classe de langue étrangère.' In Yılmaz et al. (ed.) (2004): *Linguistique Vers une Grammaire Linguistique Du Turc: Multilingual Yabancı Dil Yayınları. Tьrkiye*, pp.167-181

Vauquois, B. (1968). *A survey of formal grammars and algorithms for recognition and transformation in machine translation*, IFIP Congress-68 (Edinburgh), pp. 254-260.

Appendix A

A Course Syllabus of Massachusetts Institute of Technology

(電腦輔助翻譯譯文) 24.03 相對主義，理性和現實，2005 年春季

課程重點

本課程的特色是在四個文件的完成，作業完成的部份。

課程描述

道德標準是否絕對等於文化或者道德的框架？所謂的道德標準在所有的文化觀點看來是否都是一樣？在這個所謂的「道德」的議題之中，是否存在有：與傳統文化道德不相容、兼具或者不完全類似的思維方式？是否世界各地所有文化傳統的觀點都是一樣，所有的國家對我們的這一個世界的看法都一樣？或者其實各個地方所定義出來的道德標準，可能都存在有一些不完全相等的感受與觀點？是否目前人類科學的發展，正在邁向更接近事情最終的真相？是否科學的發展目前正以正確的方向進行？什麼又是理性？理性又應該如何定義？是否理性—可以被解釋成一個「可依據用來相信某一些概念的一個很好的理由」？是否理性-應該會相對於許多文化傳統的標準規範？另外一個議題，應該如何闡釋「自我」的涵義？是否「自我」與一個相對一致的形式相等呢？根據許多知名的學者，如孔恩，吉爾伯特哈曼，朱迪湯姆森，約翰佩里和德里克帕菲特等人，著作中的哲學理論與闡釋，我們試圖為以上這些模糊而且非常難以回答的問題，找到準確解釋的出路，本課程讓我們可以為這些人生哲學的難解的議題，開始找到一些可能正確的回答。

(Original Source Text) 24.03 Relativism, Reason, and Reality, Spring 2005, Massachusetts Institute of Technology

----

Socrates. (Image courtesy of National Institute of Environmental Health Sciences.)

課程重點

This course features four paper assignments, located in the assignments section.

課程描述

Are moral standards relative to cultures and/or moral frameworks? Are there incompatible or non-comparable ways of thinking about the world that are somehow equally good? Is science getting closer to the truth? Is rationality--the notion of a good reason to believe something--relative to cultural norms? What are selves? Is there a coherent form of relativism about the self? Guided by the writings of Thomas Kuhn, Gilbert Harman, Judith Thomson, John Perry and Derek Parfit, we attempt to make these vague questions precise, and we make a start at answering them.

(電腦輔助翻譯譯文) 教學大綱

- -

課程描述

道德標準是否是相對於所有地區的文化和傳統道德的框架？世界各地的多元文化是否存在有不兼容或不儘然類似的思維方式？是否人類的科學發展，不斷的更朝向宇宙最終事情的真相？人類科學的發展到底是否有真正的價值？什麼是理性？是否我們可以把理性解釋成：「一個可以用來相信某些觀念的一個很好的理由」？理性是否也相對於各個區域的文化標準規範？什麼又是自我？自我是否可以解釋成一個「相對一致的形式」呢？這一個課程試圖為以上的問題找到解答，本課程根據許多哲學家諸如孔恩，吉爾伯特哈曼，朱迪湯姆森，約翰佩里和德里克帕菲特等人的闡釋以及哲學理論，可能能夠為這些模糊的人生問題，開始找到一些答案。

作業

你將提交二十頁的書面作業。這些報告以及家庭作業將分成四次交出。

老師對學生修習此課程的期望

學生在寫作的時候可以自由隨意討論自己的想法與展現思考模式，當然，閱讀這個部份的作業，應該伴隨與同儕的討論互動。但是，（我應該不需要再重複地告訴你）在你交出的書面報告家庭作業的時候，必須完全為出自於你自己的真實作

品。您的書面報告不可以包括許多不必要的註解。如果你在應用某一個思想作家的想法的時候,可以運用你自己的話來重新措辭以展現你的了解程度。應該用比較簡單的英語來表達,咬文嚼字、做作的英文不會得到比較高的分數或者獎賞。只有一個途徑能夠讓您獲得分數較高的好成績:「用心琢磨思考應該要回答的問題,並制定明確的意見,用許多重點來支撐自己曾經提出來的想法。」

辯論

這一個課程一共分為三個部分。在每一部分結束的地方,我們將會舉行一個具有高水準了解度的腦力激盪辯論活動。每個人都應該要參加這一個針對課程內容討論的學術活動。一些思考上的交流可能夠幫助自己產生一些靈感與想法。

考試

HASS-D 規定,進修這一門課的同學們,至少要參與三小時的期末考試,參加考試是學生的義務,這個考試將會覆蓋整個學期老師曾經提出的教學材料。學生參加考試的時候可以帶書進場,但不能夠帶手寫的筆記,考試前老師會公佈時間和地點。本課程不會有期中考試,也就是說這門課只有期末考。

成績的分級

活動	百分率
論文	60%
期末考試	25%
課堂參與	5%
科參與和測驗	10%

上課的目標

本課程主要針對一些從來沒有接觸到哲學這一門課的學生。這一門哲學課程會為同學介紹到少數幾個相當重要的哲學議題,也就是說修這門課的學生將會學到幾個比較重要的哲學基礎學理,學生可以在這一個課程為幾個議題做更為深入研究,針對幾個比較重要的哲學議題,老師和學生將進行更有深度的探索。(另

外本課程也蜻蜓點水、兼顧廣度地，介紹許多個為各式各樣哲學的議題，雖然討論深度可能有一些不足，24.00，但從這些廣泛的議題我們可以了解的哲學學問可能存在的問題。)

(Original Source Texts) Course Syllabus of Massachusetts Institute of Technology

Course Descriptions

Are moral standards relative to cultures and/or moral frameworks? Are there incompatible or non-comparable ways of thinking about the world that are somehow equally good? Is science getting closer to the truth? Is rationality--the notion of a good reason to believe something--relative to cultural norms? What are selves? Is there a coherent form of relativism about the self? Guided by the writings of Thomas Kuhn, Gilbert Harman, Judith Thomson, John Perry and Derek Parfit, we attempt to make these vague questions precise, and we make a start at answering them.

Graded Assignments

You will submit twenty pages of written work (as required by HASS-D regulations). These will be divided among 4 papers (the HASS-D requirement is at least 3).

Expectations

Feel free to discuss the writing assignments -- and, of course, also the reading assignments -- with each other. But (as you don't need me to tell you) the written work you submit must be entirely your own. Your papers should not contain quotations masquerading as paraphrases. A statement of ideas from one of the assigned authors must be couched in your own words. Do not use any footnotes. There will be no prizes for writing in anything other than plain English. There is only one route to a good grade: thinking hard about the issues, and formulating and defending definite opinions.

Debates

The course divides into three parts. At the end of each part we will hold an in-class debate. Everyone is expected to participate. Yummy snacks are a distinct possibility.

Examination

HASS-D regulations require a three-hour final exam covering material dealt with throughout the term. The exam will be open book but not open notes, at a time and place to be announced. There will be no midterm exam.

Grading

ACTIVITIES	PERCENTAGES
Papers	60%
Final Exam	25%
Class Participation	5%
Section Participation and Quizzes	10%

Subject Matter

This course is primarily intended for students who have not had any previous exposure to philosophy. It treats a small number of important philosophical questions in some depth. (For more breadth and less depth, try 24.00, Problems of philosophy.)