



In-Depth Analysis of Machine Translation and Human Translation of Literary Book *Chinese Traditional Culture and a Community with a Shared Future for Mankind*

Xiang Long¹(✉), Ken Chen², Oluwatoba O. Bamigbade², and David L. Swenson³

¹ Guilin University of Electronic Technology and Youjiang Medical University for Nationalities, Guilin, China
longxiang@guet.edu.cn

² Guilin Normal College, Guilin, China

³ Deep Education Institute, Madison, USA

Abstract. Driven by the rapid development of AI technology and machine learning, the advancement of machine translation (MT) is constantly gaining momentum. Undoubtedly, MT has greatly improved the efficiency of translation, especially the translation of scientific and technical documents. That said, the application of MT (Google Translate in this case) has shown some limitations, particularly in literary texts. This article analyzes two English versions of the Chinese book *Chinese Traditional Culture and a Community with a Shared Future for Mankind* in lexical and syntactical aspects from a Deep Approach, points out the current challenges facing MT system, such as the translation of classical Chinese and Chinese poems, and offers some practical advice for the improvement of MT, such as the establishment of parallel corpus and translation model based on neural network. If we manage to increase the accuracy of MT in literary fields, the application of MT will surely expand enormously. It will facilitate cultural exchanges and promote cultural diversity.

Keywords: machine translation · literary texts · translation model · cultural exchanges · neural network · deep approach

1 Introduction

MT is part of a wider sphere of “pure research” in computer-based natural language processing in computational linguistics and artificial intelligence, which explore the basic mechanisms of language and mind by modelling and simulation in computer programs [1]. As information technology constantly advances, the approach and means of natural language processing are being updated. In recent years, we have entered the era of MT based on neural network, as AI is introduced into the field of MT with the support of big data and deep learning. As the quintessential examples of MT based on neural network, many translation tools and software have made their debut since 2016, such as Google Translate and SDL Trados Studio.

© The Author(s) 2023

D. Kumar et al. (Eds.): IEIT 2023, AHSSEH 10, pp. 1163–1170, 2023.

https://doi.org/10.2991/978-94-6463-230-9_139

Generally speaking, MT system is suitable for scientific and technical materials, for they tend to have repetitive and fixed features and demand accuracy and consistency. The same goes to legal texts. MT system is also capable of translating legal texts with high accuracy and smoothness, for such materials tend to contain many standard clauses and standard terms. Once the corpus is well-established with sufficient content, MT system can handle these materials like a piece of cake. Cai thinks that it's not impossible for MT system to achieve 90% accuracy with the sufficient training of neural network and learning [2].

The presence of MT system with high accuracy and efficiency has prompted a question in the translation community: Will machine translation replace human translation? To answer this question, we need to realize that there are different types of texts [3]. There is no doubt that several machine translation systems are available on the Internet that provide fine translations, but developed systems are not perfect, and their quality may decrease in some specific domains [4]. Take legal texts for instance. Despite relatively high accuracy, the translation of legal texts also requires the consideration of varied conditions in different jurisdictions, namely contexts. The translation of Chinese phrase “公益法律服务” by Google Translate into English is “public interest legal service”. Nothing is wrong with the translation result per se, but it may not be the right translation in certain contexts, if you consider the lawyers that offer such service. There are pro bono lawyers and public interest lawyers in the American legal systems, and they are quite different [5]. Thus, the lack of taking context into consideration may limit the further development of MT system in legal areas. Besides, unlike human translators, most standard modern MT systems today still work at the level of individual sentences (rather than textual level), unaware of the context.

Achieving universal translation between all human language pairs is the holy-grail of MT research [6]. Nevertheless, it is universally known that the translation of literary materials has always been their Achilles' heel. The translation of Chinese literary materials have been a challenging hurdle for MT systems and MT model programmers, because of the presence of Classical Chinese characters whose meanings vary in different contexts. The fact is that most Chinese literary materials have been done by human translators both at home and abroad. The fine translation of Chinese classics is much owed to the efforts made by Chinese scholars like Xu Yuanchong and foreign sinologists, such as Ezra Pound. Therefore, the application of MT systems in literary texts has been limited.

In this case, when we used Google Translate to translate the book *Chinese Traditional Culture and a Community with a Shared Future for Mankind* [7], many translation errors have been made. As a type of literary materials, this book features culture-loaded words and Classical Chinese characters, but most current MT systems lack the database and technical abilities to produce idiomatic and acceptable translation. In response, this paper will offer some suggestions for these existing problems. If the application of MT systems in Chinese literary texts could be substantively expanded, the translation of Chinese classics may facilitate the cultural “going out” of Chinese culture, enhance the cultural soft power of China, and promote cultural diversity [8].

2 In-Depth Analysis and Case Study

2.1 Deep Analysis

The concept of “depth” in education emerged from a variety of disciplines [9]. Current shallow education and learning practice need to be interrupted. To reach this goal, we prefer to analyze the current topic through use a deep approach. Deep approach was systemically proposed by Tochon [10, 11]. Deep approach advocates a deep integration of intradisciplinary, inter disciplinary and transdisciplinary knowledge [10, 12]. A translation must be such a deep discipline. When we mention an analysis for the case study in the context, we mean an in-depth approach.

2.2 Case Study

This part presents several examples from the book *Chinese Traditional Culture and a Community with a Shared Future for Mankind*, along with Google version and human version of translation. We analyzed the difference between the two versions in both lexical and syntactical aspects to show the accuracy of Google Translate in literary documents.

Source Text 1 (ST1): 中国讲‘和实生物’

中国	讲	和	实	生	物
zhōngguó	jiǎng	hé	shí	shēng	wù
China	value	harmony	achieve	grow	things

Google Translate (GT1): China talks about “living things in harmony”.

Human Translation (HT1): China values that “when harmony is achieved, all things will grow and thrive”.

In-depth Analysis 1: Chinese character *jiǎng*讲 as a verb has many meanings, such as “to talk about”, “to explain”, “to value”, and “to negotiate”. Obviously, Google Translate has failed to understand the right meaning of the character *jiǎng*讲 in the context, for the subject “China” is a country instead of a real person, and the verb “talk” can’t be logically placed behind inanimate subject. Other than the failure to comprehend the context, the translation doesn’t recognize the character *jiǎng*讲 here is short for the phrase *jiǎngjiù*讲究 (value, stress). This phenomenon is pretty common in Chinese language, and native speakers tend to understand the context out of question. Moreover, the character *shí*实 (achieve) as a verb has no corresponding translation in GT1, which means the MT system somehow missed that character. Also, the verb *shēng*生 (grow) has been misinterpreted by the MT system as an adjective. Out of context, the phrase *shēngwù*生物 means “(living) things”, but the character *shēng*生 is a verb in the context, meaning “to live or grow”. As a result, the translation generated by the MT system has deviated from the true meaning of ST1.

ST2: 故先王以土与金木水火杂 以成百物

故	先	王	以	土	与	金	木	水
gù	xiān	wáng	yǐ	tǔ	yǔ	jīn	mù	shuǐ
thus	ancient	King	use	earth	and	gold	wood	Water
火	杂	以	成	百	物			
huǒ	zá	yǐ	chéng	bǎi	wù			
fire	mix	to	create	hundred	things			

GT2: Therefore, the first King used soil and Jin Mu to mix fire and water, and hundreds of things.

HT2: Therefore, the ancient Kings mixed gold, wood, water, fire and earth to create all kinds of things.

In-depth Analysis 2: In GT2, the character *jīn* 金 (gold or metal) and *mù* 木 (wood) have been falsely interpreted as a name. The concept of Five Phases in Chinese philosophy is apparently excluded from the MT system database. The Five Phases as the fortune-telling system include fire, water, wood, metal or gold, and earth or soil. Moreover, the Chinese numeral *bǎi* 百 (hundred) in this context doesn't specifically refer to one hundred, but means a large number in a less specific expression. When you mix the five basic elements, there will be a myriad of combinations, thus creating a variety of things. We don't think "hundreds of" in GT2 is completely unacceptable, but there are better alternative expressions to choose, such as "various", "all kinds of", and "a variety of". Therefore, we think the common concepts in Chinese philosophy should be included in the corpus and database of MT system, so that it may recognize the Five Phases or the Five Elements in the future.

ST3: 夫仁者，己欲立而立人，己欲达而达人

夫	仁	者			
fū	rén	zhě			
man	ren	person			
己	欲	立	而	立	人
jǐ	yù	lì	ér	lì	rén
you	want	establish	conjunction	establish	people
己	欲	达	而	达	人
jǐ	yù	dá	ér	dá	rén
you	want	success	conjunction	success	people

GT3: A man of benevolence, if he wants to stand, he will establish others, if he wants to reach, he will reach others.

HT3: Now the *rén*¹ man, wishing himself to be established, sees that others are established, and, wishing himself to be successful, sees that others are successful. (Translated by A. Charles Muller).

In-depth Analysis 3: Unsurprisingly, the verb *lì* 立 has varied meanings in different contexts, such as “to stand”, “to establish”, “to set up”, and “to define”. It seems that the MT system recognized the most common and literal meaning of the verb *lì* 立, but the meaning doesn’t fit the context here. The same goes to the verb *dá* 达. There is no doubt that Classical Chinese is quite a challenge for MT system, for it requires enough existing parallel texts to establish a reliable and accurate database for access. That said, this should not be a problem anymore, since there is already a database for reference. The above sentence is quoted from *The Analects of Confucius*, and there are many English versions of this book online, such as the one translated by A. Charles Muller, whose work is available on <http://www.acmuller.net/con-dao/analects.html>. Therefore, if the parallel texts are included in the MT system, then the accuracy of the translation of well-known Chinese classics will be greatly improved.

ST4: 富者田连阡陌, 贫者无立锥之地

富	者	田	连	阡陌		
fù	zhě	tián	lián	qiānmò		
rich	person	field	connect	path		
贫	者	无	立	锥	之	地
pín	zhě	wú	lì	zhuī	zhī	dì
poor	person	no	stand	awl	of	place

GT4: The rich are in the fields, and the poor have no place to stand.

HT4: The rich own a large patch of land, while the poor have nothing at all.

In-depth Analysis 4: In GT4, the meaning of the sentence has deviated from the original one. The first half of the original text stresses the wealth of the rich, rather than their location. The second half emphasizes how destitute the poor are, since they don’t even have any space for their awl, let alone a place to sleep. Moreover, since there is a stark contrast between the poor and the rich, the conjunction “while” is better than “and” in this case. ST4 is quoted from *The Book of Han*, whose English version is also available online.

ST5: 必须抛弃冷战思维, 清除你死我活, 你衰我强, 我胜你败的非此即彼的博弈思想。

GT5: We must abandon the Cold War mentality, eliminate the all-or-nothing game thinking that you are dead, you are weak, and you are strong, and I win and you lose, and establish a fair, just, and secure state-to-state relationship mechanism.

¹ The word *rén* 仁 is perhaps the most fundamental concept in Confucian thought. It has been translated into English as “benevolence”, “altruism”, “goodness”, “humaneness”, and etc. It is a difficult concept to translate because it doesn’t really refer to any specific type of virtue or positive endowment, but refers to an inner capacity possessed by all human beings to do good, as human beings should (<http://www.acmuller.net/con-dao/analects.html>).

HT5: We must abandon the Cold War mentality, eliminate the zero-sum game mindset.

In-depth Analysis 5: The upside is that GT5 has managed to add a logical subject to the sentence, when the real subject in ST5 is omitted. Such omission is pretty common in Chinese language. However, when it comes to translating the four sets of contrastive characters: 死(die) vs. 活(live), 衰(weak) vs. 强(strong), 胜(win) vs. 败(lose), and 此(this) vs. 彼(that), GT5 is rather lengthy and redundant. Actually, they all serve the same purpose, namely referring to the win-lose, or zero-sum outcome. The word “zero-sum” is more than enough to express the meaning the writer is trying to convey. The use of “zero-sum game mindset” makes the translation very succinct and clear.

3 MT Limitations

Judging from the examples above, we have identified several major challenges facing MT systems. The first one is the translation of Classical Chinese. Classical Chinese or literary Chinese is a traditional register of written Chinese. It is based on the grammar and vocabulary of ancient Chinese, so it is very different from any modern spoken form of Chinese. Normally, the translator needs to translate Classical Chinese into modern Chinese first, then convert it into another foreign language, such as German and English. It's also challenging for MT systems to be qualified for this tough job, for they have to face the situation where one character has multiple meanings in varied contexts. If they fail to understand the original meaning, then the translation will be unacceptable.

The second challenge is that MT systems are not smart enough to apply the principle of succinctness with moderation. Currently, most MT systems are designed with the word-for-word mode, which makes the application of succinct rules rather difficult. If MT systems are granted the discretion to deploy the above principle without restriction, some key information will be missed in the translated texts.

Moreover, most standard modern MT systems today still work at the level of individual sentences, unable to distinguish different contexts. The problem is that literary text is context-based. The meaning of literary words may vary in different contexts.

4 Suggestions

In order to make MT systems qualified to translate Classical Chinese into other foreign languages, we first need to make MT systems smart enough to automatically translate from Classical Chinese into modern Chinese, because most current MT systems can understand modern Chinese easier.

To build a translation model requires the presence of a bilingual parallel corpus. The Deep Structured Semantic Alignment Model (DSSAM) is a viable method for establishing a parallel corpus by using a deep semantic matching model. The model uses the deep semantic matching model to achieve sentence alignment by synthesizing the semantics between Classical Chinese and the modern Chinese translation, and finally produces a bilingual parallel corpus [13]. Also, we need to build a MT model based on neural network for machine learning. Wang proposed a model Semi-supervision Classical Chinese Translation Neural Machine Translation (SCCT-NMT) that combines

a monolingual corpus to train the MT system to translate between Classical Chinese and modern Chinese. In a word, we need a parallel corpus to achieve the ideal state of MT systems in practice.

Apparently, the above sentence alignment model may limit the application of succinct rule in MT systems. The word-for-word translation model has a downside, namely making the translation rather lengthy and redundant. In most cases, despite the redundancy, the translation may be understandable and acceptable to target readers. Liu's survey found that 69.1% foreign participants accepted the machine translation of *The Analects of Confucius*, while the number is 92.3% among Chinese participants [14]. Nevertheless, the rapid development of AI technology may better the machine translation in the application of succinct rule. Therefore, in response to the second challenge, we may offer sufficient training for the MT systems, such as bilingual parallel corpus and human intervention. In this way, MT systems may become more and more intelligent through machine learning.

5 Conclusion

The era of AI technology is coming. MT systems will keep evolving and getting more and more intelligent. Such a highly developed MT system will be a game-changer in both the translation and IT community, despite the fact that current MT systems still require human intervention more or less, especially in literary materials. When MT systems are intelligent enough to distinguish varied contexts, then we will be arriving at a whole new phase.

The translation of Classical Chinese is a significant yet difficult task for both translators and MT system designers. At present, the translation of Classical Chinese is very dependent on human correction, but the good thing is that more and more popular Chinese classics are being translated by both Chinese and foreign scholars, such as *The Analects of Confucius* and *The Art of War*. What they have contributed is an integral part of building a more advanced MT system, for MT systems obviously need bilingual parallel corpora to learn and improve translation accuracy.

Hopefully, the MT systems will handle the current tricky problems and reveal more surprises in the future.

Acknowledgments. This study is supported by Chinese Academic Translation Project of National Social Science Fund of China “*Chinese Traditional Cultures and a Community with a Shared Future for Mankind*” (20WSHB020).

References

1. Hutchins, W.J., Somers, H.L., An Introduction to Machine Translation. San Diego, CA: Academic Press, 1992.
2. Cai, J.G., “The era of AI: will translation and interpreting major be canceled?” Chinese Science News, pp. 4 (September 4, 2019)

3. Chen, L., Long, X.: The Impact of Machine Translation on Translators' Career Development. *Journal of Guilin University of Aerospace Technology*. (3), 433-437 (2019)
4. Wołk, K., Marasek, K., Glinkowski, W.: Telemedicine as a special case of Machine translation. *Computerized Medical Imaging & Graphics*, pp. 1 (2015)
5. Chen, K., Long, X.: On expert Li Chuangshuan's papers in the field of legal English translation. *Chinese Science & Technology Translators Journal*. (02), 41-43+37 (2021)
6. Siddhant A., Bapna, A., Orhan F., et al: Towards the next 1000 languages in multilingual machine translation: exploring the synergy between supervised and self-supervised learning. *arXiv preprint arXiv: 2201.03110* (2022)
7. Zhang, L.W.: *Chinese traditional culture and a community with a shared future for mankind*. Beijing: China Renmin University Press (2018)
8. Long, X., Liu Y.M.: A deep comparison of 5 C-E translations of "While Drinking (V)": from the perspective of eco-translatology. *Journal of Xi'an International Studies University*. (01), 109-114 (2022)
9. Tochon, V.F., Busciglio, D.F. (Eds.): *Deep education across the disciplines and beyond: A 21st Century Transdisciplinary Breakthrough [C]*. Blue Mounds, WI: Deep University Press, pp. 21 (2017)
10. Tochon, F.V.: Deep Education. *Journal for Educators, Teachers and Trainers (JETT)*, (1):1-12 (2010). Retrieved from <http://www.ugr.es/%7Ejett/articulo.php?id=1>
11. Tochon, V.F.: *Help them learn a language deeply: deep approach to world languages and cultures*. Blue Mounds, WI: Deep University Press/Beijing: Foreign Language Teaching & Research Press (2014)
12. Long, X., Ju, F.A.: On the theory and practice of foreign language deep education—a review of help them learn a language deeply. *Technology Enhanced Foreign Language Education (TEFLE)*. (02), 93-96 (2020)
13. Wang, X.P.: *Research on Classical Chinese neural machine translation with limited corpus*. Harbin Institute of Technology (2018)
14. Liu, Y.T.: *Research on the acceptability of machine translation of Chinese classics: a case study of Analects of Confucius*. Guangdong: South China University of Technology (2019)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

