

On the Development and Limitation of Machine Translation

Pengcheng Geng, Xiaodong Wang*

School of Foreign Languages, Central South University, Changsha Hunan
Email: *cstdxy2008@126.com

Received: Aug. 11th, 2020; accepted: Aug. 19th, 2020; published: Aug. 26th, 2020

Abstract

We are in an age of information in which all kinds of scientific and technological products are increasingly updated. In recent years, machine translation has also set off an upsurge in the translation industry, and its advantages are gaining more and more recognition of the public. Basically speaking, machine translation has a history of about 70 years which scattered into five stages. Although it embraces just a short history, its influence is here for all to see. For example, translation tools such as Google Translation, Baidu Translation and Sogou Translation are becoming more and more popular because of their high efficiency and low cost. However, since its appearance, the question of whether machine translation can replace manual translation has been controversial [1]. This paper discusses the evolvement and shortcomings of machine translation, focusing on such limitations as mistranslation, omission and stiff translation, and also gives examples to illustrate these. And it is concluded that machine translation will not replace manual translation in the foreseeable future.

Keywords

Human Translation, Limitation, Mistranslation, Omission, Stiff Translation

机器翻译的发展与局限

耿鹏程, 王晓东*

中南大学外国语学院, 湖南 长沙
Email: *cstdxy2008@126.com

收稿日期: 2020年8月11日; 录用日期: 2020年8月19日; 发布日期: 2020年8月26日

摘要

我们所处的时代是信息时代, 各种科技产品快速更新。近几年在翻译行业也掀起机器翻译的热潮, 机器

*通讯作者。

翻译越来越多的优点受到大众的认可。总的来说, 机器翻译大致历经五个阶段, 到目前约70年的发展历史。时间虽然不久远, 但其发展之成效是有目共睹的。谷歌翻译、百度翻译、搜狗翻译等翻译工具因其具有的高效性和低成本性的优点越来越受大众偏爱。但自机器翻译入世以来, 关于机器翻译能否取代人工翻译的问题一直饱受争议[1]。本文针对机器翻译的发展历程及其存在的缺陷进行了探讨, 重点对当今机器翻译存在的误译、漏译以及表达生硬三大局限性进行了探讨, 通过对各局限性进行举例说明, 得出在可预见的未来内, 机器翻译不会取代人工翻译的结论。

关键词

人工翻译, 局限性, 误译, 漏译, 生硬表达

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

人工智能的迅速发展无疑极大的方便了人们的生活和工作, 人工智能已经渗透到我们生活的各个方面, 也包括翻译方面。早些年, 人们出国的时候必须带一名翻译人员, 现在, 一个智能翻译产品或者仅是一部智能手机即可。随着笔者对机器翻译了解的加深, 尽管被其超高的智能化所触动, 但笔者还是认为在近可预见的几年里, 机器翻译尚存不足, 不会完全取代人工翻译。笔者对机器翻译尚存的不足进行了探究, 并举例例证。

2. 机器翻译的发展

机器翻译即运用机器辅助进行的翻译。这些“机器”可以是计算机、手机、在线翻译的网页(谷歌翻译、百度翻译、搜狗翻译等)也可以是一些专门的翻译产品(比如讯飞魔飞)。

总的来看, 机器翻译的发展史可以分为五个时期: 早期探索时期(1933~1956), 第一次热潮时期(1956~1966); 商用的基于规则时期(1967~2007); 统计机器翻译时期(1993~2016); 神经网络机器翻译时期(2013 至今)。

机器翻译的研究历史可以追溯到 20 世纪三四十年代。20 世纪 30 年代初, 法国科学家 G. B.阿尔楚尼提出了用机器来进行翻译的想法。1933 年, 俄罗斯人特洛伊斯基(Peter Petrovich Troyanskii)第一次提出了机器翻译模型。1946 年, 世界上第一台现代电子数字计算机埃尼阿克(ENIAC)诞生之后, 美国科学家 Warren Weaver 提出了机器翻译的思想, 并写了《翻译》一文。1954 年, 美国 Georgetown-IBM 实验室成功利用 IBM70 计算机将 60 句俄文句子翻译成了英文。机器翻译的研究迅速铺开, 并在 1956 年后开始了持续时间近 10 年的第一个热潮。

自《翻译》发表之后, 整个机器翻译领域就分为两个流派: 语言学学派和信息论学派。语言学学派相信机器翻译需要通过对源语言和目标语言的语法分析, 解析其意义, 再进行翻译。而信息论学派则从信息论角度出发, 认为翻译是一个破解密码的过程。1966 年, 美国的自动语言处理顾问委员会发表了著名的机器翻译研究现状报告, 这份报告指出机器翻译进展缓慢, 质量糟糕, 价格昂贵, 而且看不到未来, 他们建议专注于词典开发, 这样导致的结果是美国科学家几乎有 10 年没有参与竞争。

基于规则的翻译方法的系统包括双语词典以及每种语言的语言学规则。其中还可根据类别分为直接

机器翻译、基于转换的机器翻译、中介语机器翻译、基于例子的机器翻译。1976年,以加拿大蒙特利尔大学与加拿大联邦政府翻译局联合首次开发出实用性翻译系统为标志,机器翻译步入繁荣阶段。

1993年现代统计机器翻译方法问世,基于词对齐的翻译模型。统计机器翻译崛起,但在2012年之后,统计机器翻译就逐渐进入了平淡期。

2014年,蒙特利尔大学发布了一篇关于在机器翻译中使用神经网络的论文。2016年,谷歌对外发布了一个跨越性进展—神经机器翻译。由于其质量普遍被认为已经跨过许多应用的实用性门槛,也引发了机器翻译的第二股热潮。这也是我们现在正处于的时代,市面上的神经机器翻译系统越来越多,包括国内的阿里巴巴、腾讯、百度、科大讯飞、搜狗,国外的谷歌、脸书、微软等。目前最优质的机器翻译模型也是谷歌首创的神经网络机器翻译[2]。

3. 机器翻译的优势

从机器翻译的含义和实际工作可以看出,其相对于人工翻译,使用方便,随开随用,多种语言随时切换,支持段落翻译,极其方便,只要有网就能翻译,拥有较高的便捷性和高效性[3]。对于一篇待翻译的文稿,放入计算机后,在双语语料库中进行匹配,从而找到合适的翻译。人工翻译需要几个小时才能完成的翻译量,运用机器翻译可能几秒就能完成。

面对一翻译项目,项目经理需要组成团队进行协同翻译,前后耗费时间较长,而且还需要大笔资金。在21世纪,各种运用于翻译的软件和工具层出不穷,大部分是免费的,它们不仅速度快,耗费的时间与精力也只是人工翻译的几分之一。

4. 机器翻译的局限性

利用机器翻译处理的文本大多比较简单,而且质量得不到保障。针对专业性文本,机器翻译不免乏力,容易出现误译现象。而且对于内容较多的文本,机器翻译也易出现漏译现象。对于带有感情色彩的语句,机器翻译往往达不到其表达效果,较为生硬。

4.1. 误译与漏译

误译,是指译者在翻译中因错误理解而出现的错误表达[4]。特别是文化误译,文化误译对两文化之间的正常交流有很大的影响和障碍[5]。误译现象出现的原因,一词多义,同一个词在不同语境下表现出不同意思;文化差异;专业性术语;句式结构复杂等。

① 刷脸

谷歌翻译: face brush

搜狗翻译: brush face

人工翻译: face swiping

刷脸是近几年兴起的表达,人脸生物特征识别技术的俗称,即通过人脸识别技术进行身份鉴定的过程。很明显,这里“刷脸”中的“刷”并不同于日常生活中“刷牙”的“刷”,而是“刷卡”中的“刷”。因此,不能翻译为“brush”。这是典型的不了解实际意思而进行的逐字字面翻译。

② 听说你在此次比赛中一举夺魁,恭喜你啊!

惭愧,惭愧! [6]

搜狗翻译: I heard that you won the first prize in this competition, congratulations!

Shame, shame!

谷歌翻译: I heard that you won the first prize in this competition, congratulations!

Ashamed, ashamed!

人工翻译: I heard that you won the first prize in this competition, congratulations!

Thank you very much!

在此例子中, 三个机器翻译工具的翻译大同小异, 看似符合原文, 但了解中华文化的人一眼就能看出端倪, 第二句“惭愧, 惭愧!”的翻译存在严重问题。既然是在比赛中一举夺魁, 怎么又会感到惭愧呢? 究其原因, 在于, 中西文化的差异, 在受表扬时中国人习惯以自谦的方式回复“惭愧, 惭愧!”但西方人不知道“惭愧, 惭愧!”正确的译法为“Thank you very much” [6]。

③ As in the First Symphony, there is a slow introduction, again marked Adagio molto; but while that of the First Symphony lasts a mere 12 bars, this one extends to 33 bars.

搜狗翻译: 正如在第一交响曲中, 有一个缓慢的介绍, 再次标志着慢板; 但是第一交响乐的持续时间只有 12 小节, 而这一个延伸到 33 小节。

谷歌翻译: 与第一交响乐一样, 它的演奏很慢, 再次被标记为 Adagio molto。但第一交响曲只持续 12 小节, 而这一节扩大到 33 小节。

人工翻译: 它像第一交响曲那样同样有一个缓慢的引子, 而且同样标为 Adagio molto; 但是第一交响曲的引子只有 12 小节, 而第二交响曲的引子却长达 33 小节。

通过对比发现, 机器翻译与人工翻译最大的区别在于对 introduction 的翻译。introduction 最基本的意思是“介绍”, 但是在音乐中, introduction 则表示“引子”, 而机器翻译则处理成了“介绍”, 不免让听众心生疑惑。

④ 睡得像死猪一样

搜狗翻译: sleep like a dead pig

百度翻译: sleep like a dead pig

谷歌翻译: sleep like a dead pig

人工翻译: sleep like a log

“睡得像死猪一样”是中文里的一个俗语表达, 形容人睡得香, 睡得熟。乍一看, 会让人觉得“sleep like a dead pig”直译过来就是“睡得像死猪一样”, 但这是中国人的思维。在西方文化中, 并不会将一个人睡得香与猪联系到一起, 而是木头(log), 所以“睡得像死猪一样”应该翻译为: sleep like a log.

⑤ 置身虎穴

搜狗翻译: Be in a tiger's den

百度翻译: In a tiger's Den

谷歌翻译: In the tiger's den

人工翻译: Play oneself in a lion's mouth

在中国文化里, 老虎是万兽之王, 所以就用“虎穴”形容十分凶险的地方。而在英文中, 狮子才是万兽之王, 所以在涉及到此类用“万兽之王”代指某些意思时, 要将中文的“老虎”译文“lion” [6]。

机器翻译中的漏译现象多发生于文本较长的情况。毕竟机器始终是机器, 特别是在处理较长文本的时候, 很容易出现句子的漏译现象, 这也是当下机器翻译存在的一个较大问题。

4.2. 生硬表达

表达欠佳, 这是机器翻译的一个通病。一般而言, 机器翻译只是达意, 情感和色彩是极难达到的。中华文化博大精深, 语言文字寓意丰富, 特别是对于古诗词的翻译, 机器翻译最多只能做到达意。当然, 在英译中时也会出现同样的问题。

① I love three things in this world.

Sun, moon and you.

Sun for morning, moon for night, and you forever.

搜狗翻译: 我爱这个世界上的三样东西。

太阳, 月亮和你。

太阳代表早晨, 月亮代表夜晚, 你代表永远。

百度翻译: 我喜欢这里面的三样东西世界。

太阳, 月亮和你。

太阳代表早晨, 月亮代表夜晚, 你永远。

谷歌翻译: 我爱这个世界上的三件事。

太阳, 月亮和你。

太阳代表早晨, 月亮代表夜晚, 你永远。

人工翻译: 浮世三千, 吾爱有三。

日, 月与卿。

日为朝, 月为暮, 卿为朝朝暮暮。

通过对比, 不难发现, 三种机器翻译都显得太过生硬, 表达不出真实情感。相比之下, 人工翻译就充满感情, 深情款款, 让读者能够感受到文字背后所蕴藏的情感。

② If you do not leave me, we will die together.

搜狗翻译: 如果你不离开我, 我们将一起死去。

百度翻译: 你若不离不弃, 我必生死相依。

谷歌翻译: 如果您不离开我, 我们将死在一起。

人工翻译: 问世间情为何物? 直教人生死相许。

通读一番不难发现, 三种机器翻译中, 搜狗翻译和谷歌翻译的译文太过生硬, 而且还曲解了原文。百度翻译的较为有感情色彩, 这是应当给予肯定的。相比之下, 人工翻译运用了中国文化里的经典诗句, 感情色彩更为丰富。

5. 机器翻译的前景展望

人工智能时代, 翻译技术是数字人文主义下的翻译人文和技术的融合, 是时代和语言服务行业对翻译从业者的技术要求, 是翻译生态系统的重要组成[7]。机器翻译的发展, 方便人们快速理解母语之外的语言的大致内容, 在一定程度上不用再依赖于人工翻译[8]。

但是在机器翻译取代人工翻译的路上仍然挑战重重, 就近几年来看, 机器翻译是不可能取代人工翻译的。这主要是因为对于一些翻译质量要求较高的翻译, 比如高端的商务活动、专业论文或是文学作品等的翻译, 目前机器翻译还无法胜任, 只能依靠人工翻译[9]。虽说机器翻译仍然存在较多的局限, 但其实用性与便利性肯定是不能忽视的。当前, 国内翻译产业仍然以人工翻译为主, 这是由于中国人口众多、工资水平较低、机器翻译产业的技术水平欠缺等因素的制约, 机器翻译产业的成长空间非常巨大[10]。当下的译后编辑已经慢慢崭露头角。在机器翻译处理过的基础上进行人工编辑, 无疑是将机器翻译的优势与人工翻译的优势结合起来, 将机器翻译和人工翻译的缺陷最小化。这样一来, 在机器翻译后进行人工处理, 能够将机器翻译的漏译, 误译问题解决, 并且, 还可以将生硬表达优化, 更加真实地传递原作者情感。也就是说利用机器翻译进行译后编辑不仅高速, 而且准确性高, 能够更加贴近原文。而且, 随着科技的发展, 智能化水平的提高, 机器翻译的前景也会是一片广阔!

基金项目

湖南省学位与研究生教育改革研究项目“团队化跨学科国际研究生培养模式研究”，课题编号：2019JGYB047。湖南省哲学社会科学基金项目“内地与港台地区汉英翻译能力培养比较研究”，课题编号：18JD71。中南大学深化创新创业教改项目“《中国英语能力等级量表》背景下大学英语汉英翻译能力培养目标研究”，课题编号：2019CG053。

参考文献

- [1] 张怡雪. 浅析机器翻译能否取代人工翻译[J]. 山西青年, 2020(8): 116.
- [2] 蔡强, 董冬冬. 基于 GOOGLE 神经网络汉英翻译的译后编辑研究——以科技文本为例[J]. 西南石油大学学报(社会科学版), 2020(1): 107-112.
- [3] 古丽皮娜·加迪. 人工智能背景下的机器翻译与人工翻译: 竞争与融合[J]. 国际公关, 2020(5): 176+178.
- [4] 金惠康. 跨文化交际翻译[M]. 北京: 中国对外翻译出版社, 2004.
- [5] 黄选明. 文化背景与误译[D]: [硕士学位论文]. 南宁: 广西大学, 2001.
- [6] 赵庆晶. 英语翻译中常见误译现象分析[J]. 德宏师范高等专科学校学报, 2013, 22(2): 65-70.
- [7] 王华树, 李智. 人工智能时代的翻译技术研究: 内涵、分类与趋势[J]. 外国语言与文化, 2020, 4(1): 86-95.
- [8] 尹金芳. 从科技英语翻译实例看基于神经网络的机器翻译[J]. 南昌师范学院学报, 2019, 40(6): 58-61.
- [9] 崔林艳, 虞金芳. 人工智能背景下机器翻译质量对比分析与前景展望[J]. 皖西学院学报, 2019, 35(2): 43-48.
- [10] 张星, 王建华. 意识研究视域下中国机器翻译产业发展的探源性分析[J]. 科学决策, 2019(12): 77-91.