

# 机器翻译在汉俄翻译实践中的优劣势及改进建议

吴婷婷

大连外国语大学俄语学院, 辽宁 大连

收稿日期: 2024年1月10日; 录用日期: 2024年2月22日; 发布日期: 2024年2月29日

## 摘要

近年来, 在日益增长的国际交流需求以及相关科学尤其是人工智能和计算机技术革命性的推动下, 机器翻译取得了前所未有的发展。凭借强大的技术程序和较高的任务效率, 机器翻译在很大程度上满足了人们的日常需求。然而, 由于人类语言本身的复杂性、计算机程序的不灵活性等因素, 机器翻译也存在诸多弊端。因此, 在这种情况下, 尤其是在信息和计算机技术时代, 有必要认识和了解什么是机器翻译, 并对其进行研究和改进。本论文将介绍机器翻译的定义、机器翻译的发展历史和应用, 并通过举例分析机器翻译的优缺点, 总结汉语多义词翻译的特点和难点。在文章的最后, 本文分析了机器翻译不完善的原因, 并提出了一些优化建议。

## 关键词

机器翻译, 中俄实践翻译, 利弊, 优化建议

## Advantages and Disadvantages of Machine Translation in Chinese-Russian Translation Practice and Suggestions for Improvement

Tingting Wu

School of Russian Language, Dalian University of Foreign Languages, Dalian Liaoning

Received: Jan. 10<sup>th</sup>, 2024; accepted: Feb. 22<sup>nd</sup>, 2024; published: Feb. 29<sup>th</sup>, 2024

## Abstract

In recent years, driven by the growing demand for international communication and the revolu-

tion of related sciences, especially artificial intelligence and computer technology, machine translation has achieved unprecedented development. With powerful technical procedures and high task efficiency, machine translation has largely satisfied people's daily needs. However, due to the complexity of human language itself, the inflexibility of computer programs and other factors, machine translation also has many disadvantages. Therefore, under such circumstances, especially in the era of information and computer technology, it is necessary to recognize and understand what machine translation is, and to study and improve it. This thesis will introduce the definition of machine translation, the development history and application of machine translation, analyze the advantages and disadvantages of machine translation through examples, and summarize the characteristics and difficulties of Chinese polysemy translation. At the end of this thesis, the reasons for the imperfection of machine translation are analyzed, and some optimization suggestions are put forward.

## Keywords

Machine Translation, Chinese-Russian Practical Translation, Advantages and Disadvantages, Optimization Suggestions

Copyright © 2024 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

翻译是把一种语言文字的意义用另一种语言文字表达出来的过程。翻译伴随着人类语言和社会交往的出现而产生,拥有悠久的历史。从历史发展的角度来说,翻译连接着不同的语言和文化,对于人类文化传播和文明发展有着重要意义。从语言本身的角度来说,作为人类最复杂的行为之一,翻译也是语言学的重要研究方向。

近年来,良好稳定的环境使得国际关系的迅速发展,得益于此,国家之间的交往日益密切:不仅政府和企业间的跨国合作项目越来越多,普通人也更爱去国外学习,旅游或者工作。在这样的背景下,传统的人工翻译已经不能满足越来越普遍的翻译需求。而已经有近百年技术和理论积累的机器翻译,在人工智能等新型计算机技术的加持下,获得前所未有的发展。机器翻译的产出和发展极大地改变了翻译行业,对语言学,计算机技术等科学的发展方向也产生了重要影响。目前,各类机器翻译的软件和服务层出不穷,逐渐为普通大众所接受,但机器翻译的译文质量良莠不齐的问题一直没有得到根本性解决。

出于上述原因,了解和研究机器翻译的现状和特点不仅有助于自己的专业学习,丰富自身的语言知识,对于未来就业也有重要的指导意义。

## 2. 机器翻译的发展与应用

### 2.1. 机器翻译的基本定义

机器翻译是指在计算机上完成的,将一种自然语言的文本依照其语言学特征转换成另一种语言文本的行为和结果[1]。为了进行翻译和确保文本的顺序性(连贯性),计算机中引入了一个特殊的程序,该程序运行高级翻译算法,这意味着对文本进行严格定义的操作,以便在两种语言中找到和识别相同或相似的单词或语言结构。机器翻译系统通常包括许多不同的词汇表以及提供翻译匹配所需的信息(形态、句法等)。此外,与翻译结果的分析 and 评估相关的算法工具也是现代机器翻译所必需的。

需要注意的是,尽管机器翻译技术发展迅速,但仍然没有完全自动化的机器翻译。由于相关的计算

机行业仍然不完善以及机器翻译的主要难点——自然语言的形式化差(很难被形式化) [2], 目前, 为了完成翻译过程并提高文本质量, 机器翻译仍然需要人工的帮助, 这主要体现在以下三个方面:

专业翻译人员对未来翻译过程中可能遇到的文本进行预处理(例如: 确定单词的含义); 作为必要的部分, 译者参与翻译过程以及纠正译文中的错误。

## 2.2. 机器翻译的发展简史

机器翻译与现代技术的发展和 application 息息相关, 但实际上, 机器翻译的思想可以追溯到很久以前。

17 世纪, 使用机械设备进行语言翻译的想法被首次提出, 1903 年德国科学家 W. Rieger 首先使用了“机器翻译”这个术语。现代对机器翻译的研究可以追溯到 1930 年代和 1940 年代。1933 年, 苏联科学家 P. P. 德赖弗设计了一种语言翻译机, 但由于当时技术条件差, 该机器没有投入实际运行。1946 年, 现代电子计算机 ENIAC 诞生, 这对机器翻译的出现和发展具有重要意义。1949 年, 美国学者 W. Weaver 发表了《关于翻译的备忘录》, 正式提出了机器翻译的思想, 可以说是现代机器翻译的理论开端[3]。机器翻译的第一次公开演示和应用发生在 1954 年。1954 年 6 月, 美国进行了一次机器翻译实验, 其中 60 个短语是用计算机翻译的。尽管翻译系统相当不完善, 但这个实验激发了许多专家灵感。

从那时起, 机器翻译的发展大致经历了三个阶段:

(1) 1940 年代和 1990 年代是机器翻译发展的第一个阶段, 这一时期以基于规则的翻译系统(Rule-Based)为主。这种类型的机器翻译基于目标语言和源语言的形态、句法、语义和语用模式。此类机器翻译通常包括三个不同的子系统: 逐字翻译分析系统、转移系统和语言间(跨语言学的)系统。翻译过程是这样的:

- ① 根据计算机输入词典的内容, 分析确定原文的词汇语法结构。
- ② 将原语的词汇语法结构翻译成目的语的语法结构。
- ③ 语言文本生成。

需要说明的是, 这种机器翻译虽然有很多缺点, 还远未完善, 但它很好地适应了计算机程序的特点。因此, 基于规则的机器翻译被认为是机器翻译的经典方法。

(2) 1990 年代中期~21 世纪的前十年被认为是机器翻译发展的第二个阶段。这一时期, 基于语料库的机器翻译(Corpus-Based)凭借自身的优势出现并迅速发展。基于语料库的机器翻译主要分为两种类型: 统计机器翻译和基于实例的机器翻译。统计机器翻译是使用最广泛的机器翻译形式。这种机器翻译的第一个想法是由 Warren Weaver 在 1949 年提出的。2003 年, 南加州大学教授 Joseph 使用统计模型改进了传统的机器翻译方法, 然后他发明的机器翻译系统在测试中获得了最高分[4]。与统计机器翻译一样, 基于实例的机器翻译也是一种语料库机器, 其基本思想由日本著名机器翻译长尾信提出。这种机器翻译的最大优势是与其他方法相比翻译精度更高。

(3) 目前, 随着深度学习领域研究的巨大进展, 基于人工神经网络的机器翻译(neural machine translation)逐渐兴起。这种类型的机器翻译使用大型人工神经网络。神经网络是由大量互连的简单处理器组成的系统。它的显著特点是能够自主学习同一种语言, 并在翻译过程中应用所学知识。谷歌 2016 年推出的“谷歌神经机器翻译系统”就是这种新型机器翻译的代表。

## 2.3. 当前机器翻译的实际应用

### 2.3.1. 机器翻译的种类

目前, 现实生活中的机器翻译主要分为两种:

(1) 机器翻译硬件。例如: 电子词典、翻译机、翻译笔等。此类机器翻译设备通常具有独立完整的组件, 内置翻译程序, 使用起来非常方便。最大的缺点是需要定期下载资源或购买新硬件才能更新词库。

(2) 在线翻译工具(软件)。随着互联网和手机的兴起,在线翻译工具已经成为机器翻译实际应用的重要组成部分。谷歌翻译作为最常用的翻译程序,是在线翻译工具的典型代表,在线翻译工具的优势包括:1) 翻译资源自动更新,无需花费大量时间和金钱。2) 翻译能力强。大型科技公司正在积极使用最新的翻译技术来确保在线翻译的质量。此外,很多在线翻译工具不仅提供了基本的文本翻译功能,还增加了语音或图像识别等服务,人们不再需要花费大量时间输入文本,可以快速获得翻译结果。

### 2.3.2. 使用机器翻译的场景

以前,机器翻译只被与外语或翻译密切相关的领域的人使用。例如:外语系的师生、从事跨境贸易的企业、或者直接研究机器翻译的研究人员。在国际交流日益密切和机器翻译技术不断发展的今天,机器翻译被广泛应用于公共生活的各个领域,如图书出版、国际旅游等。可以说,机器翻译已经成为一种人们文化生活的重要组成部分。

## 3. 机器翻译在汉俄多义翻译实例上的优缺点

### 3.1. 中文和俄语中多义词

多义词或歧义词是指具有两个或多个含义的词,并且这些含义之间通常存在联系。多义是语言中的常见现象之一,曾经语言学家对俄语、英语和法语进行调查,发现多义占这些语言的80%以上。“任何具有悠久文化传统的活语言中的绝大多数词都是多义的。”多义或歧义是语言发展的必然结果。一方面,随着社会的不断发展,语言应用过程中对语义表达的需求不断增长。另一方面,不同民族和民族之间的历史、文化、经济和宗教信仰的差异,也对词义的改变产生重大影响。为了语言的“经济性”,人们为了避免在认知和交际中出现大量词语造成的混乱,对一个词语赋予了多种含义[5]。总之,多义词的产生是语言系统的“有限”与社会发展变化的“无限”相互适应的结果。

作为一种语言现象,多义词的意义可以分为基本意义和外延意义两大类。前者也称为“直接意义”,一般是最常用的;后者也称为“间接意义”,可以是潜在意义的延伸,也可以是潜在意义的变化。但是,一些多义词的外延意义与主要意义无关。

作为最复杂的语言现象之一。歧义问题只能通过交互来解决[6],因为一个词可以有两个或多个含义,这极大地复杂化了工作,并可能降低文本翻译的质量。多义翻译一直是机器翻译的难题,目前还没有已彻底解决。

### 3.2. 中俄多义机器翻译研究

为了了解机器翻译的翻译过程和特点,我们对机器翻译进行了研究。本研究的目的是:多义词典型翻译错误的判定;研究机器翻译的效率和质量。

首先,我们分别选择了10个俄语和汉语中的多义词作为研究对象,并将这些多义词放入固定的句子中,然后我们使用谷歌翻译和百度翻译来翻译这些句子。最后,我们对译文进行分析,得出以下结论:

(1) 分析这些数据后,我们发现大约80%(20句中有17句)的多义翻译结果存在不同程度的错误。

(2) 翻译速度极高,翻译质量差是机器翻译最明显的特点。不到几秒钟,我们就能得到翻译后的文本,但在20个译文中,只有两个是完全准确和可以理解的。其余18句有不同程度的翻译错误。需要注意的是,俄语和汉语的一些特点,如俄语中的大小写、时态、缺乏形式顺序以及汉语中的“同音异义”等,也会影响翻译结果。因此,翻译质量差的原因是复杂的,歧义只是其中之一。

### 3.3. 机器翻译的优缺点

综合以上研究和大量相关资料,我们总结了机器翻译的优缺点。



优点是：

(1) 翻译速度快：如上所述，极快的翻译速度是机器翻译不可否认的优势。虽然翻译的质量不高，甚至个别句子出现错误。但在大多数情况下，机器翻译可以在不到两秒的时间内完成翻译过程并输出文本。对于那些对翻译质量要求不高，只需要简单理解大意的行业或人群，机器翻译可以显著提高效率。

(2) 可用性：与传统的人工翻译相比，机器翻译不受时间和空间的限制。今天，机器翻译随时可用。随着手机的兴起和互联网的发展，人们只需要一部手机就可以进行简单的文本翻译。

(3) 通用性：不断更新的计算机硬件和改进的软件算法，包括更先进的云技术，让现代计算机拥有更多的内存。巨大的计算机内存为海量语言信息的存储和更新提供了可靠的物质基础，而这种语言信息不仅可以包括词汇、语法等语料库知识，还可以包括专业(经济、法律、农业等)词典。对于人工翻译来说是难以想象和困难的。

(4) 低成本：从综合和长期来看，机器翻译的综合成本比人工翻译的成本低 40%以上。以国内几家大型翻译公司为例：百度翻译、有道翻译等，平均每千字翻译价格为 300 元，考虑到文本范围、期限等，人工翻译的实际成本要高得多[7]。相比人工翻译，从长远来看，机器翻译要便宜得多，更何况很多在线翻译工具都是免费的。

(5) 保密性：一些翻译文本，尤其是与商业和政治有关的文本，是属于机密的，而机器翻译的一个重要优点是它的保密性——机器翻译的整个过程是由计算机上的程序进行的，这意味着输入的信息不会被泄露。机器翻译的保密性大大降低了翻译过程中关键信息泄露的风险，避免了不必要的损失。

如今，机器翻译的大规模使用证明了其无可比拟的优势和广阔的发展前景。但同时，机器翻译的弊端也不容忽视。

机器翻译的劣势主要体现在以下两个方面：

(1) 翻译质量差。

翻译质量差是机器翻译问世以来一直存在的问题，也是阻碍机器翻译进一步发展的主要障碍。2016 年，Google 推出了基于神经网络的翻译系统，同年，在文章(*A Neural Network for Machine Translation, at Production Scale*)中，Google 表示相比传统的人工翻译和其他类型的机器翻译错误，神经基于网络的翻译系统下降了 60%，这个翻译数据超过了所有现有的已发布结果。

机器翻译的主要错误体现在：选择的词或短语的含义与原文的上下文不符；语法使用不正确，词汇规则导致句子中单词和短语的顺序错误；有些词没有翻译，等等。

(2) 输入原文和输出文本的类型和形式受到严格限制。

目前，绝大多数机器翻译系统只能“理解”和翻译计算机系统程序可以识别的文本或单词。虽然一些机器翻译可以从声音或图像中提取信息，但它们效率不高，有时提取的信息不正确。因此，人们不得不花时间打字，这在一定程度上阻碍了机器翻译的发展。另一方面，机器翻译后的翻译文本主要以文字和声音的形式展示，并不能完全满足人们的需求。相反，人工翻译更加灵活，可以处理 PPT 格式文件、视频等不同类型的翻译项目。

## 4. 机器翻译优化建议

在上一节中，我们简要概述了影响机器翻译发展的两个主要问题。为了让评论更有针对性，我们首先需要了解导致问题的原因。

### 4.1. 机器翻译存在弱点的原因

从计算机技术的角度来看：

(1) 计算机程序工作的特殊性决定了它与人类语言之间存在着一种自然的而无法解决的矛盾。计算机

程序具有确定性、有序性和有穷性的特点，它们必须按照一定的规则和逻辑运行。翻译时，计算机只能按照预先设计好的程序来识别和翻译文本，一旦文本超出自己的语料库，计算机就无法理解，翻译的结果就不正确。而人类语言是“灵活的”和“多变的”，我们不能按照严格的计算机程序来使用语言。在大多数情况下，如果词汇没有按照既定的语法进行排列，甚至拼写错误，人们仍然可以正确理解其含义。

(2) 计算机技术还不够完善。在翻译领域，即使是最先进的翻译系统——神经机器翻译也无法与人脑匹敌。例如：人脑可以在多义的各种含义中准确地判断出最适合上下文的含义，并且可以自动调整句子的顺序，理解句子的意思。中国数学家、语言学家周海忠曾在《机器翻译五十年》一文中指出：如果人们还没有理解大脑是如何对语言进行模糊识别和逻辑判断的，那么机器翻译做到“真实、准确、优美”(信、达、雅)是不可能的。

从语言的角度来看：

(1) 每种语言都有自己的特点，不同语言之间存在很大差异，几乎没有完全对等的词汇。例如：在从俄语翻译成其他语言的过程中，由于复杂的格系统、时态和缺乏稳定的顺序，造成了大量的问题和错误。翻译难度很大。

(2) 单词或句子中的文化内涵和人们使用语言时的情绪是机器翻译目前无法理解和翻译的东西。

(3) 机器翻译的质量还取决于原文的风格和题材。研究表明，源文件的风格越正式，翻译效果越好。任意文学文本的机器翻译几乎总是质量不令人满意。

#### 4.2. 机器翻译优化建议

(1) 未来机器翻译的类型应该以基于统计和神经的机器翻译为主。我们不应将翻译“限制”在使用既定的语言规则编写人机程序中，而应基于海量统计数据 and 机器学习来构建翻译所需的语料库。人类语言是“活的”并且不断发展，在实际使用中，我们的语言在很多情况下不受绝对规则的限制。因此，要更好地翻译语言，我们只能依靠统计和机器学习。

(2) 要建立成熟完善的翻译质量反映机制。首先，我们可以邀请用户对机器翻译的质量进行评价，并奖励高质量评价和纠正。其次，我们对翻译过程进行统计，重点关注经常出现的单词、句子和文本类型，以及翻译结果被接受的频率。最后，根据上述数据得到的结果，我们在翻译过程中显示 2 个或更多的结果供用户选择。然后你需要重复前两个步骤，不断调整翻译结果，使其尽可能准确，以满足用户的要求。

(3) 改进计算机技术，使机器翻译能够接收和输出更多形式的文本，以满足人们的需要。如前所述，原始输入和输出文本的类型和形式受到严格限制。目前，机器翻译无法应对视频、图像、各种文件格式等的翻译问题，这是机器翻译相对于人工翻译的主要缺点之一。因此，我们还需要解决机器翻译系统与其他计算机系统的交互问题：图像识别系统、语音识别系统、视频信息处理系统等。像今天的物联网一样，机器翻译系统必须是整个系统的重要组成部分，能够接收其他部分提供的数据，并以图像和声音的形式共享翻译结果。而这对于今天的计算机发展来说仍然是一个巨大的问题。

总的来说，机器翻译优化不能只关注翻译质量问题，机器翻译本身存在的其他问题也需要我们大量的时间来解决。

### 5. 结语

机器翻译具有人工翻译无法比拟的优势，随着信息技术的不断发展和国际交流的日益密切，使用计算机程序执行翻译任务是不可阻挡的发展趋势。尤其是近年来人工智能和神经网络翻译技术的改进和应用，为机器翻译的发展提供了新的思路 and 空间。未来，机器翻译甚至可能在翻译任务中完全取代人类，这也是很多科学家的最高愿望。

同时需要强调的是，机器翻译的问题非常复杂，对其研究还不完善。目前，机器翻译的结果只能用于“表面审阅”，机器翻译辅助工具还需要进一步完善，以适应未来发展的需要。因此，完全独立于人类的机器翻译，在今天乃至未来几十年都未必会出现。

综上所述，本文介绍了机器翻译的类型和发展历史、优缺点，希望能为外语学生的学习和就业提供新思路，也希望同学们能顺应潮流，更好地运用计算机工具，提高翻译效率和质量。

## 参考文献

- [1] 史芸. 人工智能翻译的知识伦理与知识建构[J]. 当代外语研究, 2023(6): 156-165.
- [2] 余红卫. 翻译工程师需求与翻译人才技术能力培养研究[J]. 管理工程师, 2023, 28(6): 59-63.
- [3] 袁洋, 杨柳. 机器翻译在对外交流中的应用及效果评估[J]. 百科知识, 2023(33): 60-61.
- [4] 郝立涛, 于振生. 基于人工智能的自然语言处理技术的发展与应用[J]. 黑龙江科学, 2023, 14(22): 124-126.
- [5] 许宗瑞. “欧洲翻译技术硕士”培养: 模式、特色与启示[J]. 外语界, 2023(6): 56-63.
- [6] 吴尹清, 李德俊. 翻译错误研究: 问题与对策[J]. 外语教学, 2023, 44(6): 79-86.
- [7] 王鼎, 王汝发. 自然语言处理中机器翻译发展态势的哲学思考[J]. 绥化学院学报, 2023, 43(11): 78-80.