

知识丛书

机器翻译浅说

刘 湧 泉
高 祖 舜 著
刘 倬



1973-1974 年第 1 期

机 器 制 造 技 术

第 1 卷 第 1 期
1973 年 1 月
第 1 页



机 器 翻 譯 淺 說

刘湧泉 高祖舜 刘 倬著

《知識丛书》編輯委员会編

一九六四年·北京

知識就是力量。一个革命干部需要有古今中外的丰富知識作为从事工作和学习理論的基础。《知識丛书》就是为了滿足这个需要而編印的；内容包括哲学、社会科学、自然科学、历史、地理、国际問題、文学、艺术和日常生活等知識。为了使这一套丛书編写得更好，我們期望讀者們和作者們予以支持和合作，提供意見和批評。

《知識丛书》編輯委员会

机器翻譯淺說

刘湧泉 高祖舜 刘 倬著

科学普及出版社出版

(北京市西直門外鄭家灣)

北京市书刊出版业营业许可证出字第 112 号

北京市通县印刷厂印刷 新华书店发行

开本 787×960 1/32 印张 6 7/8 字数 100,000

1964 年 8 月第 1 版

1964 年 8 月北京第 1 次印刷

印数 8,820 定价 0.55 元

总号 058 统一书号 13061·028

目 次

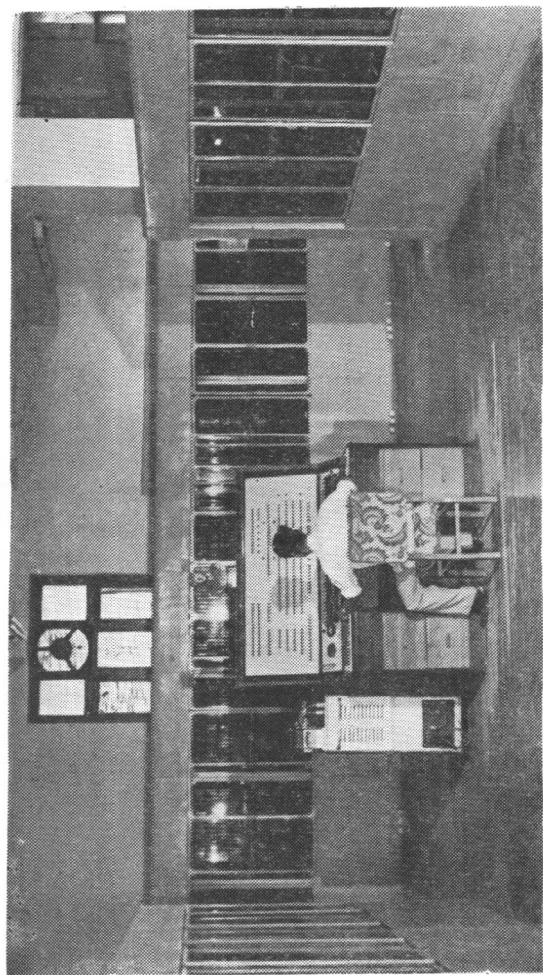
一	从唐僧談起	5
二	人与机器	11
	电子计算机的构造和工作原理	12
	語言和代碼	14
	語言代碼的輸入和存儲	24
	机器如何“思維”	36
	譯文輸出	39
三	語言的可譯性	42
四	机器詞典	53
	机器詞典是实现机器翻譯的基础	54
	机器詞典和一般詞典	56
	机器詞典的內容	61
	机器詞典的工作	72
	俄汉机器詞典	84
五	机器語法	98
	“詞对詞”翻譯和“句对句”翻譯	101
	机器对語法提出的三点要求	107
	俄汉机器語法	115
六	从俄語到漢語	140
七	机器会听和会說外国話	173

八	机器翻譯类型和媒介語	179
九	未来的万能翻譯博士	189
	語言研究中的問題	190
	数学家的任务	198
	工程师的工作	199
十	机器翻譯簡史	202
附录一	机器翻譯試驗略举	209
附录二	俄汉机器翻譯規則系統略語、符号 查对表	211

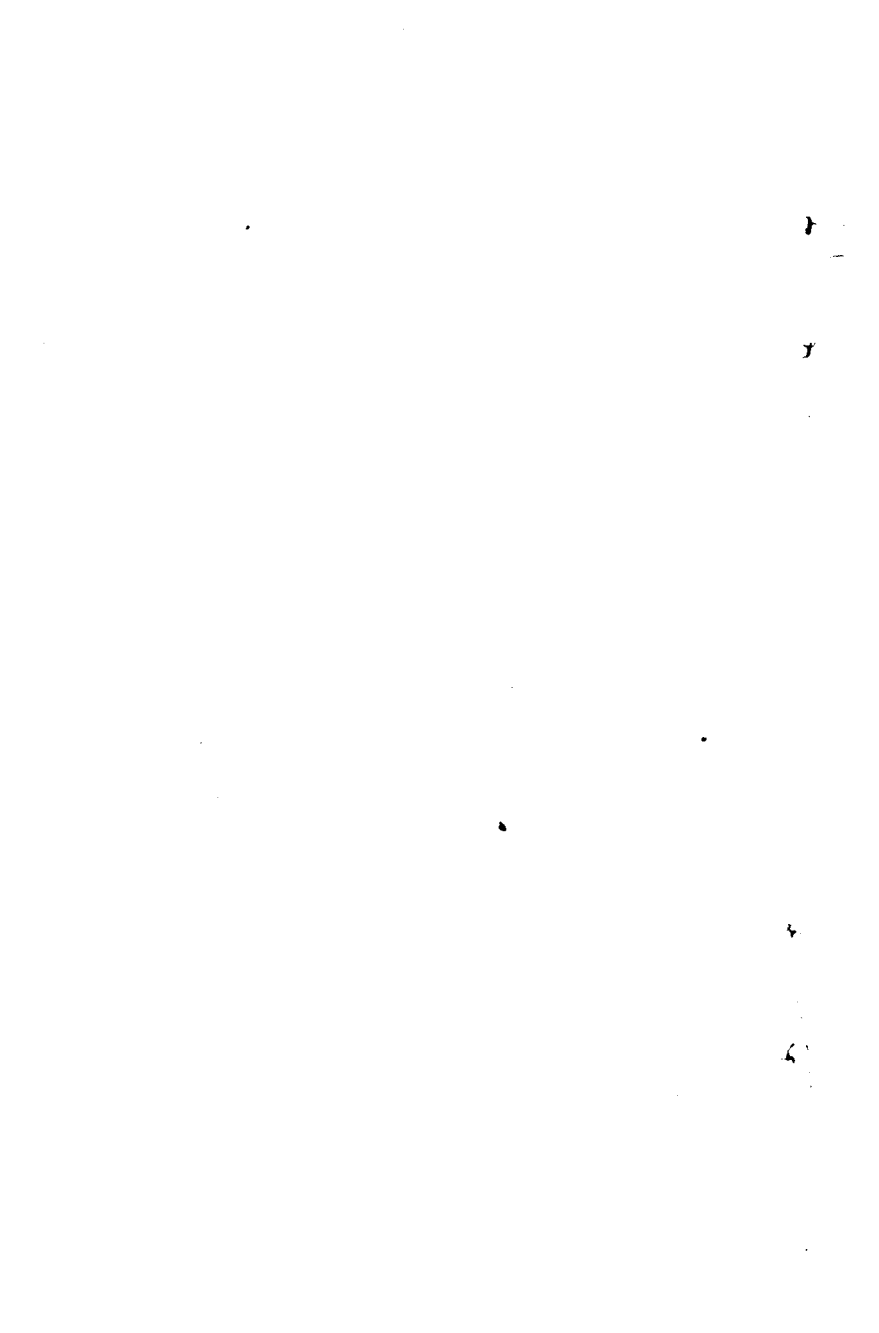
目 次

一	从唐僧談起	5
二	人与机器	11
	电子計算机的构造和工作原理	12
	語言和代碼	14
	語言代碼的輸入和存儲	24
	机器如何“思維”	36
	譯文輸出	39
三	語言的可譯性	42
四	机器詞典	53
	机器詞典是实现机器翻譯的基础	54
	机器詞典和一般詞典	56
	机器詞典的內容	61
	机器詞典的工作	72
	俄汉机器詞典	84
五	机器語法	98
	“詞对詞”翻譯和“句对句”翻譯	101
	机器对語法提出的三点要求	107
	俄汉机器語法	115
六	从俄語到漢語	140
七	机器会听和会說外国話	173

八	机器翻譯类型和媒介語	179
九	未来的万能翻譯博士	189
	語言研究中的問題	190
	数学家的任务	198
	工程师的工作	199
十	机器翻譯簡史	202
附录一	机器翻譯試驗略举	209
附录二	俄汉机器翻譯規則系統略語、符号 查对表	211



电子计算机全图



一 从唐僧谈起

不同国家或不同民族，一般都有自己独特的语言文字。因此，不同国家或不同民族之间的交际，必须通过翻译进行。翻译是桥梁，这话一点也不夸张。只要有国际或族际交往出现，便一定有翻译问题存在。没有翻译，正如没有桥梁一样，大河两岸的人民就只能望洋兴叹了。

翻译的历史非常悠久。我国在两三千年以前的周朝就有翻译官了(那时叫“象胥”)。周秦之后，文化交流的需要增多，翻译工作才日益发展起来(主要是翻译佛经)。到了唐代，出现了一个伟大的翻译家玄奘(俗称唐僧)。他终身从事翻译工作，译出经卷数量之多，质量之精，可以说是前无古人了。他不仅从事梵译汉的工作，而且还从事汉译梵的工作(把老子的《道德经》译成了梵文)。明清以来，翻译重点从佛经转到科学著作。晚清至五四运动以前，随着人们对社会改革要求的增长，社会政治作品和文学作品的翻译逐渐发达起来。五四运动至中华人民共和国成立的三十年间，翻译工作

無論在內容或形式上都有了很大變化。就內容而言，馬克思列寧主義經典著作開始翻譯過來，許多國家的文學名著陸續介紹過來，自然科學著作也有了較多的譯本。就形式而言，白話文在譯本中已經佔據了統治地位。新中國成立以後，翻譯工作進入了一個新階段：語種顯著增多，並且不只是把外國東西介紹進來，而且還將我國的東西介紹出去（例如，黨和政府的一些重要文件現在已用20種文字出版，又如《人民畫報》經常用17種文字出版）。翻譯作品的種類也大大為增加。數量較以前有成倍的增長，質量也有極大的提高。

隨着文化的發展和彼此交往的頻繁，翻譯工作日益增加。現在已經增加到這樣一個地步：需要和可能之間發生了很大矛盾。一方面需要翻譯大量的東西，而另一方面却由於翻譯人材的限制不能滿足這種需要。這個矛盾，隨着時間的推進變得越來越尖銳了。

那麼，為什麼翻譯和翻譯需要之間的矛盾會如此尖銳起來呢？原來是，需要翻譯的材料在不斷增加，而翻譯的技術卻變動不大。就拿目前世界上出版的情況來看，每年世界各國出版的書籍有25萬種，其中科學技術書籍5萬餘種；各國出版的定期刊物已超出8萬種，其

中，自然科学和技术科学的刊物将近4万5千种，每年刊载300多万篇论文。可是翻译技术的改进情况怎么样呢？我们知道，唐僧之后，直到如今，中间经过一千三百多年，翻译技术上还没有发生很大变化。

为了解决这个矛盾，人们曾想到两条出路：第一，统一语言文字——这就是各种类型的国际语的设计；第二，提高翻译效率——这就是翻译工作机械化和自动化的理想。第一条路的问题不是我们讨论的对象，这里就不详谈了。这里只准备谈谈有关第二条出路——翻译机械化和自动化的问题。

人们想利用机器来做翻译的念头很早就出现了。1933年，有个叫特罗扬斯基的苏联人，便曾设想用一种机械装置进行翻译工作。但是，由于当时技术条件不够，他的理想终于落了空。直到1946年，由于电子数字计算机的问世，这才给翻译自动化创造了技术条件。电子数字计算机是二十世纪科学技术中最伟大的成就之一。这种机器的运算速度是惊人的，它能在一秒钟之内完成上万次，乃至数十万次的运算。因此，如果把语言转化成数字代码并加以程序化，就有可能利用这种神速的机器，来代替人做翻译，从而大大地提高翻译的速度。

翻譯自动化的彻底实现，将标志着翻譯史上的一次大革命。

利用电子计算机做翻譯的最初研究，是在1946年英美学者中間进行的。1952年，在美国麻省理工学院召开了第一次机器翻譯會議。1954年初，在紐約国际商业机器公司举行了第一次机器翻譯試驗。其后，立刻引起了世界各国語言学家、数学家、物理学家、技术工程师、邏輯学家等多方面学者的注意，对机器翻譯中的有关問題进行了广泛的研究和深入的討論，并且在机器上进行了一系列不同語言之間的翻譯試驗，終于使机器翻譯成为一个独立的專門学科。

机器翻譯是在語言学、数学和計算技术这三門科学的基础上建立起来的一門边缘学科。机器翻譯利用了这三門科学的成就，同时反过来也促进了它們的发展。

从語言学方面来看，机器翻譯要利用严密方法对語言的形式和結構进行分析和描写，建立語言之間形式的和結構的对应規律，并在此基础上制定一套能在机器上按一定程序自动把一种語言轉換成另一种語言的規則系統。为此，机器翻譯利用了传统語言学关于語言分析和描写的基本理論。但是，由于传统語言学所建立的語言理論，往往是以邏輯語义为出发点的(如关

于“詞类”、“句子成分”、“句子”等等的概念和定义就是如此)；所以，远远不是传统語言学的全部理論都能用在机器翻譯上(因为机器是不懂意义的)。机器翻譯只能利用传统語言学关于語言形式—結構分析的理論。以上談的只是机器翻譯和語言学的关系的一个方面。另一方面，机器翻譯也能够反过来促进語言理論的进一步丰富和发展。比如說，机器翻譯的出現要求語言学家研究制定形式化和严密化的語言分析方法，机器翻譯的試驗可以客觀地检查語言研究的成果，因此，在一定意义上讲，它已成为驗證語言分析描写理論的严密性和科学性的場地。

在数学方面，机器翻譯首先是利用数理邏輯中的算法理論来建立翻譯規則系統；进而又依靠計算数学中的程序設計来实现翻譯規則系統。机器翻譯的发展，要求数学家解决翻譯程序設計自动化的問題，要求数学家和語言学家建立更密切的关系，利用数学工具——集論、概率論、数理統計、格論、数理邏輯——来分析和描写語言，以促进数理語言学的建立和发展。

在計算技术方面，机器翻譯利用了各种电子計算机順利地进行了試驗。但是，一般的电子計算机对机器翻譯來說具有很大的局限性。例如，現代电子計算机还没有完善的文字直接

輸入輸出裝置，信息存儲能力太小，运算系統过于复杂等。因此要求計算技术工作者，研制一种專門适合翻譯工作的电子翻譯机。

机器翻譯与其他許多方面（如外語教学，机器自动查詢，机器自动作文摘等等）也有密切关系，它的研究方法和研究成果，能直接或間接地在这些方面得到应用。

机器翻譯的实现，不但能使人类从一部分繁重的脑力劳动中解放出来，而且对于国民經济和科学文化的发展具有极其重要的意义。利用机器翻譯可以迅速地、及时地获得外国科学技术的情报，掌握外国科学技术的发展动向。

目前，机器翻譯还处在研究試驗的阶段，离实际应用还有一段距离。但是，随着科学技术的长足发展，依靠各方面专家的一致努力，机器翻譯进入人类文化生活已經不在遙远的未来，而是指日可待了。

机器究竟是怎样做翻譯的呢？語言又是怎样变成代碼，并在机器中存儲和运算的呢？机器在进行翻譯时，要不要象人那样查看詞典和掌握語法呢？机器在翻一句話时要分哪些步驟呢？机器翻譯中有哪些难题和机器翻譯将来能达到什么水平呢？在本书的下面几节里，讀者将找到这一系列饶有趣味的問題的答案。

二 人与机器

大家知道，翻譯是一种相当复杂的思維活动。为了把某一种外語譯成漢語，首先必須熟悉这种外語，掌握这种外語的詞匯和語法。仅仅背熟外語的單詞和弄懂外語的語法还不够，要进行翻譯還必須弄清外語和漢語在詞匯和語法上的对应关系，也就是說要弄清外語和漢語在表达方式上的异同。否則，翻譯工作就不能順利进行。比方說，俄語的 *язык* 和英語的 *tongue* 都有“舌头”和“語言”的意思，而在漢語里就找不到一个既能表示“舌头”，又能表示“語言”的詞。（不过，在某些特殊情況下，古漢語中“舌”也可作“語言”解；如“舌人”、“舌耕”等）。所以，在翻譯的時候，需要根据原文的上下文，決定是譯成漢語的“舌头”，或是譯成“語言”。決不能隨意地拿“舌头”去翻譯俄語（或英語）中不同場合出現的 *язык*（或 *tongue*），也不能拿“語言”去譯俄語（或英語）中不同場合出現的 *язык*（或 *tongue*）；不然會鬧出很多笑話來。又如我們說：“中國的各族人民”，俄語和