

吴会芹

华北水利水电学院外语系 河南 郑州 450045

论著选摘

【摘要】如今,机器翻译的研究已展现勃勃生机,各种翻译软件也应运而生:言的各种矛盾,机器实现全自动语言翻译的目的尚待时日。机器翻译只适用于某规则性语言。

【关键词】机器翻译;自动翻译;人工调整;自然语言

1 机译的起源

自古以来,人们就萌发了利用机器进行翻译的梦想。古希腊时代,人们试图设计一种理想化的语言叫自然语。17世纪,笛卡尔(Descartes)和莱布尼兹(Leibniz)试图在统一的数字代码基础上编写字典,提出用机器词典来克服语言障碍的设想。17世纪中叶展开的“普遍语言”的运动,旨在运用逻辑的基础上,创造出一种无歧义的语言。其中最著名的成果便是维尔金斯(John Wilkins)在《关于正式符号和哲学语言的论文》(An Essay towards a Real Character and Philosophical Language, 1668)中提出的中介语(Interlangua)。中介语的设计试图将世界上所有的概念和试题加以分类和编码,有规律地排列概念和实体,并根据它们各自的特点和性质,给予不同的记号和名称。

二十世纪初,德国学者里格(W. Rieger)提出了一种数字语法(ziffern grammatik),这种语法加上词典的辅助,可以利用机器将一种语言翻译成另一种语言。此,“机器翻译”这一术语(德语ein meehanisches Uebersetzen)首次被使用。

二十世纪三十年代初,法国工程师阿尔楚尼(G. c. Artsouni)提出用机器进行语言翻译的想法,并在1933年7月22日获得了一项“机械脑”的专利。该“机械脑”翻译机可以记录火车时刻表和银行帐户,尤其适合被用作机器翻译。后来,因二战的爆发,其安装最终夭折,但是,它的展出却引起科学界的轰动。

真正对机器翻译进行研究应该说是从布恩(AD. Booth)和韦弗开始的。他们研究的是自动词典(automatic dictionary),从1954年1月7日公开研制计算机开始,机器翻译进入一个繁荣发展的时期。从那时起,很多国家都投入了大量的人力、物力从事这方面的研究和开发。

2 机译的发展

随着科学技术的发展,机器翻译已经取得了很大进展。进入九十年代后,随着电脑的不断升级换代,机器翻译的文字处理能力也随之增强。在发达国家,机器翻译不仅用于文字处理系统,而且正在朝着智能声控翻译通信技术的方向发展。其中一些具有特殊用途或专用的系统,如电话订票系统,称国际上最为成熟的机器翻译系统。近来,我国机器翻译的研究也发展很快。几年来许多公司都推出了高科技机器翻译软件,如天津通译、中软通译软件等。最近,南京月亮公司、实达铭泰公司、雅信译霸公司都先后推出汉化软件翻译产品。中国科学院计算所等单位联合开发研制了“863智能翻译系统”。该系统实现了对原文的语法、语义和常识的一体化分析,具有面向多语种翻译软件环境。我国台湾地区的“功学电脑自动翻译系统”,可在“通讯”、“历史法律”、“军事”等方面进行计算机翻译。机器翻译的发展如雨后春笋展现出勃勃生机。

3 机译的主要问题

虽然机器翻译的研究和发展速度很快,翻译质量也有了很大提高,但是,毋庸置疑,机器翻译仍然存在着很多问题。其中质量问题是机器翻译的主要问题。有些翻译软件能够把结构简单的句子及常用会话翻译得很标准,可是遇到与汉语结构迥异的句子,就会产生问题。如

原文: I don't know the boy who is waiting outside.

机器译文: 我不认识那个男孩,谁在外边等着。

人工译文: 在外边等着的那个男孩我不认识。

Matrix英汉翻译系统是史晓东(glandel Shi)于1990-1995期间编写完成的(罗选民等, 1999)。该系统翻译质量有了很大改善,它存在处理固定词组中规定了一些固定词组,同时在判断语义歧义时运用复合特征集,如catch up, catch up with属于固定词组,但又规定catch后按vehicle译成“赶上”。因此,在翻译“Nidia can not catch the bus.”时,Matdx译文为“Nadia不能赶上公共汽车”。然而,该系统虽然能够对一些常用词组进行翻译,但对一些相对复杂的成语、习语却未能得以正确分析。如:

原文: He always asks for moon.

Matrix译文: 他总是正在请求月亮。

传统的机器翻译是符号处理(Symbolic)系统。它所刻画的偏重于语言能力(Language Competence)而非语言行为(Language Performance)。因此系统很难应付一些非规则语言现象,例如,错误的起始(false start),不合语法的语句等等。目前,从已推出的实用机器翻译系统的译文看,对文一般不会有太大问题。但对一些稍长的句子或结构较复杂的句子,译文质量就不能令人满意,有时简直是不可卒读(冯志伟,1999)。计算机翻译词(n,)+动词”或“v.+v.+11.”这一解释,因为它和英语句子的句法根本不相容。但在可能接受的句法模型范围内(如“n.+v.+adv.”,“+adv”,“n.+、n.”,“修饰语+n.+v.”等等),计算机就能够选择了。

一些非规则性语言如果通过人工翻译,译者可以从语篇知识、历史背景、文化差异甚至非语言因素(如身势语)中得到准确的翻译。但是,由于矛盾,机器翻译很难达到这个标准。

为解决这一问题,笔者认为机译研究需要更多的语言学研究者的参与。在进一步深入研究规则语言的同时,加强对非规则性语言的研究。使工作范围和工作效率都有大幅度的提高。

4.3 增加语用因素参数

机器翻译计划中所用的模式通常是以成份结构语法的某一形式——“预示分析法”(Predictive analysis)、“从属关系语法”(dependency grammar)、“投射语法”(projective grammar)或“后进先出存储语法”(pushdown store grammar)等等为基础的。但这些语法业已表明,原则上是不适用于自然语言。有时是会产生歧义,如:“I saw a girl with a telescope at the table.”,对于“at the table”所限定的先行词就会产生歧义。虽然语言学家在分析自然语言时并未考虑到它会产生歧义,但在计算机翻译中,这种歧义随时都有可能产生。利用人工翻译,可以根据原著及各种信息对译文的歧义进行解释,但是,机器翻译却难以顾及到这么多因素。就目前的技术看,歧义仍是机器翻译的一大难题。

对歧义的排除主要是依靠语用因素来进行的。因此机译设计中也应该考虑到以语用因素做参数,“虽然对语用类型的划分的可行性有多高目前朝这个方向做出的任何努力都不会白费。

4.4 扩大对不同语体的语言规律的研究

从目前看,多数成功的机器翻译是比较程式化的文字,如天气预报等。而用机器翻译文学作品则比较困难,用机器翻译诗歌等文学语言更为困难。从某种意义上说,机器地译出一首诗就是创作另一首诗(马修斯,引自Wilss,1989)。“……按定义,诗是不可译的,只能进行创造性的换位(雅可布森,引自wilss,1989)。

传统的机器翻译需要大量的人力和专家知识用于设计中间语言和编写语法,而这个过程又是枯燥乏味的。更大的问题是,设计出来的中间语言只用于一个特定领域。一旦更换了机器翻译的问题领域,整个设计过程又得从头再来。

计算机只能凭借其有限的内部规则——大量的文字、句子进行迅速处理,而缺乏对自然语言的语义、语境的理解,因此,译文总会有这样或那样的问题。机器翻译之后人的后续工作仍然很大(陈光火,1999)。此外,对一些复杂的句子的分析依然很困难,往往遭到失败,不过,正是由这种相对的失败起了推动作用,使机器翻译学认清了自身的发展方向。

5 结语

机器翻译的目标是通过计算机实现不同自然语言的自动翻译,以解决人类社会的语言障碍。然而,由于语言与机器的矛盾,想要用计算机把丰富的语言写得淋漓尽致,惟妙惟肖尚待时日。难怪有人指出:用程序法重新组织表层话语结构,推出尽善尽美的译文,使翻译程序达到标准化,是一种乌托邦。人类对自然语言的认识是无尽的,对机器翻译的研究也是无尽的。但不管怎么说,机器翻译的质量将会越来越好,越来越成熟。以机器翻译为核心的产业,将是21世纪最有生气的先导产业,而高性能机器翻译无疑又是这一先导产业之中最为重要的高尖科技领域(章国英,1995)。因此,我们要满怀信心迎接挑战。

机器翻译的发展也将对英语教学和研究产生深刻的影响。首先,机器翻译的发展将促进语言学科和理工学科的融合。如上所述,机器翻译的很多问题是靠IT学科所能解决的,它需要语带来许多挑战和机遇。第二,随着机器翻译质量的提高和电脑的普及,与机器翻译相关的课程,如“机译原理”,将成为外语专业学生的必修课程。二十一世纪的合格译员,必然是既能熟练掌握口笔译技能,又能熟练操作、应用机译软件的新型人才。第三,机译的发展将对外语学科中许多门类的教学与研究产生推动作用。机译需要建立庞大的语料库,这将推动词汇学、词典学的发展,同时也为这些门类的教学提供丰富的素材。机译需要研究双语的结构异同,这将促进对比语言学在更深、更精的层次上发展。此外,语法学、文体学、术语学和普通翻译理论本身都会从中面临挑战、发挥作用,并籍此得到自身的发展和飞跃。

参考文献略

文章选自《外语电化教学》2003(4)

 [返回主页](#)



版权信息:

本主页版权所有: 北京师范大学现代教育技术研究所; 管理员信箱: ysqetc@21cn.com; 电话: 010-62206922。要获取最佳浏览效果, 请使用800*600分辨率模式。