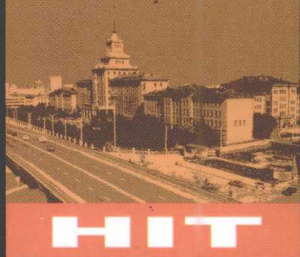


Acquire Translation Knowledge Based on Phrase Evaluation



计算机·数学系列

基于短语评价的翻译知识获取

张春祥 高雪瑶 著



哈尔滨工业大学出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS



计算机·数学系列

Acquire Translation Knowledge Based on Phrase Evaluation

基于短语评价的翻译知识获取

● 张春祥 高雪瑶 著



哈尔滨工业大学出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

内 容 提 要

本书主要包括:机器翻译概述、基于语言学知识的汉-英短语对齐、基于评价的短语翻译对优化、基于错误驱动的翻译知识获取和基于译文评价的翻译知识优化。本书介绍了机器翻译的研究背景、理论基础和算法描述,并给出了相应的实验结果。本书是自然语言处理中翻译知识获取方面的专著,反映了作者近年来在这一领域中的主要研究成果。

本书具有内容新颖、结构清晰、语言简练和实验完整的特点,可作为大专院校及科研院所从事自然语言处理、机器翻译和知识获取研究的高年级本科生、研究生的教材和参考书籍。此外,本书也可以作为相关领域的教师、科研人员和从事机器翻译系统开发的工程技术人员的参考书。

图书在版编目(CIP)数据

基于短语评价的翻译知识获取/张春祥,高雪瑶著. —哈尔滨:哈尔滨工业大学出版社,2012.2

ISBN 978-7-5603-3504-9

I. ①基… II. ①张… ②高… III. ①词对词翻译
(机器翻译)-研究 IV. ①H085

中国版本图书馆 CIP 数据核字(2012)第 015550 号

策划编辑 刘培杰 张永芹

责任编辑 刘 瑶

出版发行 哈尔滨工业大学出版社

社 址 哈尔滨市南岗区复华四道街 10 号 邮编 150006

传 真 0451-86414749

网 址 <http://hitpress.hit.edu.cn>

印 刷 黑龙江省教育厅印刷厂

开 本 787mm×1092mm 1/16 印张 14.25 字数 263 千字

版 次 2012 年 2 月第 1 版 2012 年 2 月第 1 次印刷

书 号 ISBN 978-7-5603-3504-9

定 价 48.00 元

(如因印装质量问题影响阅读,我社负责调换)

前言

1947年,美国洛克菲勒基金会自然科学部主任华伦·韦弗提出把计算机用于语言翻译,机器翻译的概念由此产生。在此后的几十年中,自然语言处理技术和统计学方法的不断完善使机器翻译得到了充足发展。翻译知识的规模和质量是决定自动翻译成败的关键性因素。目前,如何利用自然语言处理技术、语言学知识和机器学习算法从大规模语料中挖掘翻译知识,已成为制约系统快速发展的瓶颈问题。因此,研究翻译知识自动获取方法,对于机器翻译及自然语言处理将具有重要的理论意义和现实应用价值。

本书提出了一套翻译知识自动获取及优化方案。在知识获取方面,探索了平行双语语料的短语对齐技术,并研究了从短语翻译对中抽取模板的方法;在知识优化方面,探索了短语翻译对的评价及过滤技术,同时研究了基于译文评价的模板自动优化方法。在此基础上,实现了一系列的翻译知识自动获取及优化工具,将获取的模板用于机器翻译系统中,以提高译文输出质量。

第1章是概述,主要介绍了机器翻译的语言学资源和相关技术。

第2章研究了汉-英双语语料的短语对齐方法。针对汉语和英语语法体系异构的问题,综合利用统计方法和语言学知识

来实现双语句对的短语对齐;讨论了短语翻译对的定义、短语对齐的几种策略以及对齐过程中的相关问题;同时,使用“分析-分析-匹配”策略、译文等价树方法和译文中心语块扩展方法来实现短语对齐,并提出了消解策略对译文边界冲突进行处理。

第3章针对自动获取的短语翻译对中存在大量噪声的问题,提出了源语言-目标语双语评价特征来对短语翻译对进行评分。评价特征主要包括:短语译文直译率、短语翻译概率和源语言-目标语短语长度差异。使用评价特征对自动获取的汉-英短语翻译对进行评价过滤。在评价特征的基础之上,综合使用多种机器学习算法来提高评价及过滤的性能。

第4章在短语翻译对的基础之上,使用机器学习方法来获取翻译模板,提出了基于错误驱动的翻译模板自动获取方法;使用汉语的词法特征、词性特征、句法特征、核心结点信息和机器翻译词典从短语翻译对中提取学习实例,并利用错误驱动的学习方法从中获取模板。

第5章提出了一种新的翻译模板评价机制。该机制不仅考虑了模板对训练语料库译文评测分数的总体贡献,而且也考虑了模板正确翻译语言现象的能力,因而可以更准确地对其进行评价。本章分别给出了这种评价机制在逐步删除策略与逐步添加策略中的具体数学表达形式。

本书受国家自然科学基金“基于中心扩展对齐的汉-英统计机器翻译研究”,编号:60903082 资助。

第3章、第4章和第5章由哈尔滨理工大学的张春祥完成。第1章和第2章由哈尔滨理工大学的高雪瑶完成。张春祥负责了本书的统稿和协调工作。

由于作者水平有限,书中不足之处在所难免,希望读者批评指正。

张春祥 高雪瑶
2011年10月于哈尔滨



哈尔滨工业大学出版社刘培杰数学工作室

已出版(即将出版)图书目录



书 名	出 版 时 间	定 价	编 号
新编中学数学解题方法全书(高中版)上卷	2007—09	38.00	7
新编中学数学解题方法全书(高中版)中卷	2007—09	48.00	8
新编中学数学解题方法全书(高中版)下卷(一)	2007—09	42.00	17
新编中学数学解题方法全书(高中版)下卷(二)	2007—09	38.00	18
新编中学数学解题方法全书(高中版)下卷(三)	2010—06	58.00	73
新编中学数学解题方法全书(初中版)上卷	2008—01	28.00	29
新编中学数学解题方法全书(初中版)中卷	2010—07	38.00	75
新编平面解析几何解题方法全书(专题讲座卷)	2010—01	18.00	61

数学眼光透视	2008—01	38.00	24
数学思想领悟	2008—01	38.00	25
数学应用展观	2008—01	38.00	26
数学建模导引	2008—01	28.00	23
数学方法溯源	2008—01	38.00	27
数学史话览胜	2008—01	28.00	28

从毕达哥拉斯到怀尔斯	2007—10	48.00	9
从迪利克雷到维斯卡尔迪	2008—01	48.00	21
从哥德巴赫到陈景润	2008—05	98.00	35
从庞加莱到佩雷尔曼	2011—08	138.00	136
从比勃巴赫到德·布朗斯	即将出版		

数学解题中的物理方法	2011—06	28.00	114
数学解题的特殊方法	2011—06	48.00	115
中学数学计算技巧	2012—01	48.00	116
中学数学证明方法	2012—01	58.00	117
数学趣题巧解	2012—03	28.00	128

数学奥林匹克与数学文化(第一辑)	2006—05	48.00	4
数学奥林匹克与数学文化(第二辑)(竞赛卷)	2008—01	48.00	19
数学奥林匹克与数学文化(第二辑)(文化卷)	2008—07	58.00	36
数学奥林匹克与数学文化(第三辑)(竞赛卷)	2010—01	48.00	59
数学奥林匹克与数学文化(第四辑)(竞赛卷)	2011—08	58.00	87



哈尔滨工业大学出版社刘培杰数学工作室

已出版(即将出版)图书目录



书 名	出版 时间	定 价	编 号
发展空间想象力	2010-01	38.00	57
走向国际数学奥林匹克的平面几何试题诠释(上、下)(第2版)	2010-02	98.00	63,64
平面几何证明方法全书	2007-08	35.00	1
平面几何证明方法全书习题解答(第2版)	2006-12	18.00	10
最新世界各国数学奥林匹克中的平面几何试题	2007-09	38.00	14
数学竞赛平面几何典型题及新题解	2010-07	48.00	74
初等数学复习及研究(平面几何)	2008-09	58.00	38
初等数学复习及研究(立体几何)	2010-06	38.00	71
初等数学复习及研究(平面几何)习题解答	2009-01	48.00	42
世界著名平面几何经典著作钩沉——几何作图专题卷(上)	2009-06	48.00	49
世界著名平面几何经典著作钩沉——几何作图专题卷(下)	2011-01	88.00	80
世界著名平面几何经典著作钩沉(民国平面几何老课本)	2011-03	38.00	113
世界著名数论经典著作钩沉(算术卷)	2012-01	28.00	125
世界著名数学经典著作钩沉——立体几何卷	2011-02	28.00	88
世界著名三角学经典著作钩沉(平面三角卷Ⅰ)	2010-06	28.00	69
世界著名三角学经典著作钩沉(平面三角卷Ⅱ)	2011-01	28.00	78
世界著名初等数论经典著作钩沉(理论和实用算术卷)	2011-07	38.00	126
几何学教程(平面几何卷)	2011-03	68.00	90
几何学教程(立体几何卷)	2011-07	68.00	130
几何变换与几何证题	2010-06	88.00	70
几何瑰宝——平面几何500名题暨1000条定理(上、下)	2010-07	138.00	76,77
三角形的五心	2009-06	28.00	51
俄罗斯平面几何问题集	2009-08	88.00	55
俄罗斯平面几何5000题	2011-03	58.00	89
计算方法与几何证题	2011-06	28.00	129
463个俄罗斯几何老问题	2012-01	28.00	152
近代欧氏几何学	2012-2	38.00	162



哈尔滨工业大学出版社刘培杰数学工作室 已出版(即将出版)图书目录



书 名	出 版 时 间	定 价	编 号
超越吉米多维奇——数列的极限	2009-11	48.00	58
初等数论难题集(第一卷)	2009-05	68.00	44
初等数论难题集(第二卷)(上、下)	2011-02	128.00	82,83
谈谈素数	2011-03	18.00	91
平方和	2011-03	18.00	92
数论概貌	2011-03	18.00	93
代数数论	2011-03	48.00	94
初等数论的知识与问题	2011-02	28.00	95
超越数论基础	2011-03	28.00	96
数论初等教程	2011-03	28.00	97
数论基础	2011-03	18.00	98
数论入门	2011-03	38.00	99
解析数论引论	2011-03	48.00	100
基础数论	2011-03	28.00	101
超越数	2011-03	18.00	109
三角和方法	2011-03	18.00	112
谈谈不定方程	2011-05	28.00	119
整数论	2011-05	38.00	120
初等数论 100 例	2011-05	18.00	122
最新世界各国数学奥林匹克中的初等数论试题(上、下)	2012-01	138.00	144,145
算术探索	2011-12	158.00	148
初等数论(Ⅰ)	2012-01	18.00	156
初等数论(Ⅱ)	2012-01	18.00	157
初等数论(Ⅲ)	2012-01	28.00	158
组合数学浅谈	2012-02	18.00	159
同余理论	2012-02	38.00	163

哈尔滨工业大学出版社刘培杰数学工作室

已出版(即将出版)图书目录

书 名	出 版 时 间	定 价	编 号
历届 IMO 试题集(1959—2005)	2006—05	58.00	5
历届 CMO 试题集	2008—09	28.00	40
历届国际大学生数学竞赛试题集(1994—2010)	2012—01	28.00	143
全国大学生数学夏令营数学竞赛试题及解答	2007—03	28.00	15
历届美国大学生数学竞赛试题集	2009—03	88.00	43
前苏联大学生数学竞赛试题及解答(上)	2012—03	28.00	169
前苏联大学生数学竞赛试题及解答(下)	2012—03	38.00	170
整函数	2012—1		161
俄罗斯函数问题集	2011—03	38.00	103
俄罗斯组合分析问题集	2011—01	48.00	79
博弈论精粹	2008—03	58.00	30
多项式和无理数	2008—01	68.00	22
模糊数据统计学	2008—03	48.00	31
受控理论与解析不等式	2012—03		165
解析不等式新论	2009—06	68.00	48
反问题的计算方法及应用	2011—11	28.00	147
建立不等式的方法	2011—03	98.00	104
数学奥林匹克不等式研究	2009—08	68.00	56
不等式研究(第二辑)	2012—02	68.00	153
初等数学研究(I)	2008—09	68.00	37
初等数学研究(II)(上、下)	2009—05	118.00	46, 47
中国初等数学研究 2009 卷(第 1 辑)	2009—05	20.00	45
中国初等数学研究 2010 卷(第 2 辑)	2010—05	30.00	68
中国初等数学研究 2011 卷(第 3 辑)	2011—07	60.00	127
数阵及其应用	2012—02	28.00	164
不等式的秘密(第一卷)	2012—02	28.00	154
初等不等式的证明方法	2010—06	38.00	123
数学奥林匹克不等式散论	2010—06	38.00	124
数学奥林匹克不等式欣赏	2011—09	38.00	138
数学奥林匹克超级题库(初中卷上)	2010—01	58.00	66
数学奥林匹克不等式证明方法和技巧(上、下)	2011—08	158.00	134, 135



哈尔滨工业大学出版社刘培杰数学工作室

已出版(即将出版)图书目录



书 名	出版 时间	定 价	编 号
500 个最新世界著名数学智力趣题	2008-06	48.00	3
400 个最新世界著名数学最值问题	2008-09	48.00	36
500 个世界著名数学征解问题	2009-06	48.00	52
400 个中国最佳初等数学征解老问题	2010-01	48.00	60
500 个俄罗斯数学经典老题	2011-01	28.00	81

数学 我爱你	2008-01	28.00	20
精神的圣徒 别样的人生——60 位中国数学家成长的历程	2008-09	48.00	39
数学史概论	2009-06	78.00	50
斐波那契数列	2010-02	28.00	65
数学拼盘和斐波那契魔方	2010-07	38.00	72
斐波那契数列欣赏	2011-01	28.00	160
数学的创造	2011-02	48.00	85
数学中的美	2011-02	38.00	84

最新全国及各省市高考数学试卷解法研究及点拨评析	2009-02	38.00	41
高考数学的理论与实践	2009-08	38.00	53
中考数学专题总复习	2007-04	28.00	6
向量法巧解数学高考题	2009-08	28.00	54
新编中学数学解题方法全书(高考复习卷)	2010-01	48.00	67
新编中学数学解题方法全书(高考真题卷)	2010-01	38.00	62
新编中学数学解题方法全书(高考精华卷)	2011-03	68.00	118
高考数学核心题型解题方法与技巧	2010-01	28.00	86
数学解题——靠数学思想给力(上)	2011-07	38.00	131
数学解题——靠数学思想给力(中)	2011-07	48.00	132
数学解题——靠数学思想给力(下)	2011-07	38.00	133
2011 年全国及各省市高考数学试题审题要津与解法研究	2011-10	48.00	139
新课标高考数学——五年试题分章详解(2007~2011)(上、下)	2011-10	78.00	140,141
30 分钟拿下高考数学选择题、填空题	2012-01	48.00	146
高考数学压轴题解题诀窍(上)	2012-02	78.00	166
高考数学压轴题解题诀窍(下)	2012-03	28.00	167
300 个日本高考数学题	2012-03		142



哈尔滨工业大学出版社刘培杰数学工作室 已出版(即将出版)图书目录



书 名	出版 时间	定 价	编 号
中等数学英语阅读文选	2006-12	38.00	13
统计学专业英语	2007-03	28.00	16
方程式论	2011-03	38.00	105
初级方程式论	2011-03	28.00	106
Galois 理论	2011-03	18.00	107
代数方程的根式解及伽罗瓦理论	2011-03	28.00	108
线性偏微分方程讲义	2011-03	18.00	110
N 体问题的周期解	2011-03	28.00	111
代数方程式论	2011-05	28.00	121
动力系统的不变量与函数方程	2011-07	48.00	137
基于短语评价的翻译知识获取	2012-02	48.00	168
闵嗣鹤文集	2011-03	98.00	102
吴从斡数学活动三十年(1951~1980)	2010-07	99.00	32
吴振奎高等数学解题真经(概率统计卷)	2012-01	38.00	149
吴振奎高等数学解题真经(微积分卷)	2012-01	68.00	150
吴振奎高等数学解题真经(线性代数卷)	2012-01	58.00	151
钱昌本教你快乐学数学(上)	2011-12	48.00	155

联系地址:哈尔滨市南岗区复华四道街 10 号 哈尔滨工业大学出版社刘培杰数学工作室

网 址:<http://lpj.hit.edu.cn/>

邮 编:150006

联系电话:0451-86281378 13904613167

E-mail:lpj1378@yahoo.com.cn

◎ 目 录

第 1 章 概述	1
1.1 机器翻译的发展历程	3
1.2 语料库语言学	4
1.2.1 语料库	5
1.2.2 平行双语语料	7
1.3 词典资源	10
1.4 机器翻译方法	12
1.4.1 基于规则的机器翻译	13
1.4.2 基于语料库的机器翻译	14
1.4.3 基于混合策略的机器翻译	24
1.5 词义消歧	25
1.5.1 词义消歧的方法	25
1.5.2 词义消歧的应用	27
1.6 译文评价	29
1.6.1 译文评价的发展历史	31
1.6.2 译文评价方法	32
1.7 翻译知识获取的意义	34
1.8 本章小结	36
参考文献	36
第 2 章 基于语言学知识的汉-英短语对齐	39
2.1 词汇对齐研究	40
2.1.1 统计词汇对齐	41

2.1.2	基于语言知识的词汇对齐	45
2.2	短语对齐	48
2.2.1	短语翻译对的定义	49
2.2.2	短语翻译对获取	50
2.2.3	汉-英句法分析树的结构对齐	55
2.2.4	基于译文等价树的短语翻译对获取	73
2.2.5	基于译文中心扩展的汉-英短语翻译对获取	76
2.2.6	源语言短语译文边界冲突消解	81
2.3	实验及数据分析	83
2.4	本章小结	87
	参考文献	87
第3章	基于评价的短语翻译对优化	91
3.1	机器学习方法	93
3.2	短语翻译对评价特征	96
3.2.1	短语译文直译率	96
3.2.2	短语翻译概率	99
3.2.3	源语言-目标语言短语长度差异	103
3.3	短语翻译对分类器	111
3.3.1	感知器模型	112
3.3.2	支持向量机模型	115
3.3.3	多重线性回归模型	119
3.3.4	主动学习	123
3.4	实验及数据分析	124
3.5	本章小结	135
	参考文献	135
第4章	基于错误驱动的翻译知识获取	137
4.1	翻译模板的定义	139
4.2	基于模板的翻译模型	142
4.3	学习实例	148
4.4	翻译模板获取	153
4.4.1	错误驱动学习	154
4.4.2	基于错误驱动的翻译模板获取	155
4.5	翻译模板的组织	164
4.6	实验及数据分析	168
4.7	本章小结	175

参考文献	175
第 5 章 基于译文评价的翻译知识优化	177
5.1 自动译文评价技术	179
5.1.1 n-gram 模型	180
5.1.2 Bleu 译文评测方法	181
5.1.3 Nist 译文评测方法	182
5.2 基于译文评价的翻译模板库自动优化	183
5.2.1 翻译模板库自动优化框架	184
5.2.2 使用频度信息过滤初始模板库	185
5.2.3 翻译模板的评价机制	186
5.2.4 翻译模板优化算法	189
5.2.5 N 重交叉优化	191
5.3 翻译知识的领域适应	192
5.4 实验结果及数据分析	197
5.4.1 翻译模板频度过滤	197
5.4.2 θ 值的确定	198
5.4.3 优化性能的比较	199
5.4.4 机器翻译系统领域自适应	205
5.5 本章小结	211
参考文献	211

概述

第

1

章

随着计算机技术的不断发展以及互联网的日益普及,现代社会进入了信息时代。人们所面临的问题不再是信息匮乏,而是浩如烟海的不同语种信息。单纯依靠人工方式来实现不同语言之间的翻译已无法满足人们的需求,因而人们必须使用计算机来自动地将一种语言转换为另一种语言,让计算机来完成繁重的翻译处理任务。从计算机诞生之日起,实现人类语言的自动翻译就一直是人工智能研究中的重要课题。机器翻译(Machine Translation, MT)是通过计算机来实现不同语言文本之间的相互转化。简单地说,机器翻译就是用计算机将一种自然语言翻译为另一种自然语言。通常,机器翻译是指在没有人参与的情况下,由计算机来自动地实现文本转换。

在翻译过程中,如果需要人来完成部分转换任务或者需要人对转换结果进行校对,或者需要计算机提供相关信息来帮助翻译者进行翻译,这种情况被称为计算机辅助翻译(Computer Aided Translation, CAT)。与全自动机器翻译相比,计算机辅助翻译系统是一种人机交互系统。在这种翻译模式中,计算机负责辅助翻译人员执行转换任务,不但给翻译人员提供一些词汇、术语和短语的翻译知识,而且从已翻译过的文本中,查找相同或相似语句的译文,使翻译人员避免不必要的重复劳动,进

而提高翻译工作的效率。其实现原理是:在翻译记忆库和翻译实例库中,搜索相同或相似的句子和短语,并给出参考译文。计算机辅助翻译技术特别适合科技专著、科技文献、产品说明书、使用手册和联合国文件等篇幅较长并且语言重复现象较多的文本翻译。这种技术能使翻译人员充分地利用已有的翻译资源,避免重复翻译劳动。在这种翻译方式下,翻译人员只需要专注于新内容的翻译。

机器翻译是自然语言处理领域中的一个重要分支,它与计算语言学、人工智能、统计学习和自然语言理解之间存在着密不可分的关系。从计算机诞生之日起,人们就开始尝试使用计算机来对语言现象进行表示和处理。自然语言处理的研究,最初就是从机器翻译开始的。随着信息时代的到来以及因特网的迅猛发展,“信息爆炸”已经成为信息处理领域中的瓶颈问题,不同语种之间的信息交流更加大了这一问题的严重性。不同语言之间的翻译工作越来越迫切,并且工作量也越来越大。如何利用计算机高效率的信息处理能力去突破不同语种之间的障碍,就成为全人类所面临的共同课题。机器翻译是解决这个问题的有力手段,近半个世纪以来,机器翻译技术的研究一直在曲折中前进。尽管经历了几十年的发展,形成了众多理论、方法和技术,但是,现有系统的译文输出质量仍然不能令人满意,仍有诸多难题需要解决。

由于自然语言现象本身的复杂性,目前,机器翻译仍然是最困难的决策问题之一。自然语言的复杂性表现在它既是灵活多变的,同时又具有严格的一面。一个句子可以有多种不同的翻译方式,可以直译,也可以意译,同时,一种意思也可以有多种不同的表达方法。一种语言往往又有其特定的语法规律和表达习惯,这就是长期以来形成的约定俗成的规则,而且,这种规律和习惯往往难以用简单的逻辑来加以概括。因此,自然语言处理本身就是一个复杂的问题。而机器翻译需要同时对两种或者多种语言进行处理,这就变得难上加难了。很多语言在语法规律和表达习惯上差别很大。例如,像英语这样的曲折语往往强调语法规则,句子中的各个成分的位置也比较固定,而像汉语这样的非曲折语,其表达方式呈现出多样化的态势,各个成分在句子中的位置也更加灵活。

此外,语言之间的差别还体现在句子所提供的信息量的不对称性。例如,英语中的大部分句子都有相应的时态,而汉语却没有。在翻译时,就需要根据时间、助词和上下文等信息来推断译文应该采用哪种时态。在翻译过程中,如果缺乏这些信息,其译文输出质量就难以保障。

1.1 机器翻译的发展历程

1947年,美国洛克菲勒基金会自然科学部主任华伦韦弗提出把现代计算机用于语言翻译,机器翻译的概念由此产生。

在最初十几年中,机器翻译得到了迅猛的发展。据不完全统计,当时有20多个国家的70多个机构在从事机器翻译的研究工作。1950年,英国研制了第一台用于自动翻译的机器。在早期的机器翻译系统中,所使用的技术基本上都以双语词典为基础来进行直接翻译,同时,又附加了一些简单的语法分析与生成规则。在直接翻译方法中,源语言分析和目标语言生成没有明显的界限。这种方法主要依靠完备的翻译词典和词法分析,然后使用文本处理程序把源语言文本转换成目标语言文本。直接翻译方法的效果与人们的要求还相差很远。

1966年,ALPAC的报告给机器翻译研究做出了否定的结论,宣称:在目前给机器翻译以大力的支持还没有多少理由,其研究遇到了难以克服的语义障碍。在ALPAC报告的影响下,许多国家的机器翻译研究陷入了低潮。

20世纪70年代初期,计算语言学、人工智能和计算机技术取得了长足的发展,给机器翻译研究带来了希望和曙光。在20世纪70年代后期,机器翻译研究开始复苏,相继出现了一些实用化的机器翻译系统,如美国的SYSTRAN系统、加拿大的TAUM-METEO系统、日本日立公司的ATLAS系统和法国编织研究所的TITUS-IV系统。1975年以后,我国的机器翻译研究也进入了恢复和发展的新时期。在这一时期中,计算语言学的进展使得机器翻译方法不再局限于直接翻译。此时,出现了更多的间接翻译方法,如基于中间语言的翻译方法和基于转换的翻译方法等。

20世纪80年代以来,机器翻译研究受到了各国的高度重视,许多发达国家相继投入巨额资金来开展全国性和多国性的联合研究,其中包括:欧共体的EUROTRA、日本的Mu和美国CMU机器翻译研究所。此时,基于间接翻译的方法得到了长足发展,转换分析的层次从简单的语法层次发展到了语法和语义相结合的更深层次。同时,还出现了基于实例的翻译技术和基于统计的翻译技术。

1980年以后,我国也研制了几个机器翻译系统,包括黑龙江大学设计的由LISP语言支持的机器翻译专用语言NT-83和RI-84,中国人民解放军军事科学院研制的科译1号。

20世纪90年代以后,中科院计算所开展了基于人工智能的新一代机器翻译系统研发工作,并成功实现了一个面向通用领域的高性能英-汉翻译系统