

* 科学家 *

我国模式识别领域的开拓者

——记模式识别与知识工程专家戴汝为

郝 宏

(自动化研究所)

1992 年 6 月底,在巴基斯坦名城伊斯兰堡附近美丽的风景区参加第 17 届 Nathiagali Summer College 讲习班的世界各国的科学家,正在聆听一位来自中国的学者的报告。这位学者就是中国科学院学部委员、自动化研究所研究员戴汝为,报告的题目是“人工神经网络及其在手写体汉字中之应用”。

他说运用人工神经网络研制手写体汉字的识别系统,是从思维科学和心理学角度,以人的模式识别为参照,着眼于模拟形象思维,参考了记忆表象产生的认知过程,大脑皮层中存在检测器、知觉与感觉的关系,辩认是数据驱动与概念驱动的加工过程等观点,吸收和利用了模式的语义、句法描述等模式识别方法中的精华,构成一个具有学习和联想功能的非线性动力学系统。目前它的部分技术指标已达到国际领先水平,这一成果将进一步促进信息处理技术的发展。戴汝为的这个报告,在与会者中引起了强烈反响,巴基斯坦原子能委员会的专家对这一成果非常感兴趣,希望能把它运用到“乌尔都”语的识别方面(乌尔都语属阿拉伯语系)。为此,当地电视台、广播电台进行了大量报道。

戴汝为,1932 年 12 月 31 日出生在云南昆明。少年时正值抗战时期,就读于西南联大附小、附中,受到了良好的基础教育。1951 年他以优异的成绩考取了清华大学数学系,1952 年因院系调整,又就读于北京大学数学力学系。1955 年大学毕业后,即分配在力学研究所搞工程控制论,戴汝为曾有幸得到刚回国的著名科学家钱学森的直接指导。由于善于洞察国际学术研究的新动向,他逐渐形成了自己独特的、力求选择最新、最前沿课题的学术思想。经过多年的辛勤努力,他在模式识别、知识工程、思维科学等领域取得了丰硕的研究成果,在国内同行中享有很高的威望。

60 年代,最优控制的研究在国际上处于高潮。苏联科学家提出了著名的极大值原理,但这一原理在处理具体问题时存在相当大的计算困难,戴汝为和他的同事们用两年时间,解决了用极大值原理解最速控制、终值控制的数值计算问题,其结果具有重要的理论意义与实际应用价值。这是他在难度很大的科研领域上第一次崭露头角。

从 1970 年开始,戴汝为率先在国内开展模式识别的理论研究工作。20 多年来,大量的科研成果使他逐步成为我国该学科领域的杰出代表之一。由他领导的课题组,1972 年取得了“打印体字符(英文)的相关匹配识别系统”成果,1974 年取得了“手写数字的自动识别(实验样机)”成果,1976 年至 1980 年,又连续在《国外科学》和《自动化》杂志发表研究结果,系统地阐述了他们在模式识别理论研究上的贡献。

80年代初,戴汝为做为访问学者,在美国 PURDUE 大学与著名模式识别专家、美国工程科学院院士、台湾科学院院士傅京孙教授合作,在具有实际意义的“文法推断”工作上提出了一类文法的推断算法,更重要的是把统计模式识别与句法模式识别有机地结合起来,提出了新的语义、句法模式识别方法。这种方法简单、合理,从某种意义上讲,更符合人的思维实际,开辟了一个把模式识别和思维科学联系起来的新的科研领域。这一成果获中科院 1986 年科技进步奖二等奖,并受到钱学森教授的充分肯定与好评,钱老曾表示:“我深感戴汝为同志是我国模式识别方面的权威之一,学术造诣深”。

近年来,在原有基础上,戴汝为又深入研究了模式识别中的模式结构,描述与知识工程中知识表达的内在联系以及人机交互与智能接口的有关问题,把属性文法做了进一步开拓与推广,并在国内率先开展了人工神经网络用于知识工程的研究。用人工神经网络通过学习来进行模式识别、联想记忆与形象思维,提供了模式识别与知识表达的统一模型。1989 年出版了他的《人机交互与人工神经网络在专家系统中的应用》论文集,对这一新的研究领域具有指导意义,为设计知识系统提供了强有力的工具。这一时期他主持完成的“人工神经网络软件包‘Neunet’”,为亚运会工程建设做出了贡献,并获 1990 年中国科学院科技进步奖二等奖。

在知识工程方面,戴汝为冲破传统思想的束缚,采用人工神经网络解决知识获取问题,并以模式分析为出发点,把模式描述与知识表达统一起来,进而把人工智能中传统的符号处理与联接机制(神经网络)结合起来,提出利用物理符号处理、定性物理、可视知识及人工神经网络等模型来综合知识系统的学术思想,并在这一思想指导下完成了核反应堆故障诊断、中医妇科病诊断等几个新型专家系统。近 30 年来,他在国际权威性杂志及国内一级学报上发表论文 80 余篇,出版专辑两集,成为国际模式识别学术界的知名学者。

1992 年 6 月 10 日至 13 日,在深圳举行的“深圳 92 计算机高技术成果与产品展览会”上,戴汝为指导研制的“超级汉字识别系统”参加了展览。该产品包括印刷体汉字识别系统、联机手写体汉字识别系统、脱机特定人及非特定人识别系统和手写体数字识别系统,实用前景广阔。一些香港和深圳的公司主动要求合作开发,甚至于当场购买。展览会后,在“863”主题办公室主办的“全国汉字识别与语音识别评比会”上,该系统经过专家组三天的严格测试,脱机特定人手写体汉字识别的识别率和印刷体汉字识别的文本识别均取得第一名。

戴汝为非常注重对青年学生的培养。1985 年他被国务院学位委员会特批为“模式识别与智能控制”专业博士生导师,至今已培养了 30 余名硕士生、博士生。1987 年他在一次科技展览会上看到了刘迎健的工作成果,认为刘是位有才华的青年,应该及时深造培养。于是经他多方努力,终于把刘迎健破格录取为自己的在职博士生。几年来,在他的悉心指导下,刘迎健完成了 400 万字的手写体汉字样本库。这是当今国际上最大的手写体汉字样本库。在此基础上完成的“手写体汉字识别系统的理论与实践”的研究成果,获 1992 年度中国科学院自然科学奖一等奖,所研制的系统体现了语义-句法方法及人工神经网络方法的优点,在与日本、美国、台湾、新加坡等同项工作的竞争中走在前列,具有十分明显的中国特点。日本模式识别专家、神户大学田中荣一教授参观后称赞道:“我参观后印象最深的是你们的手写汉字识别系统。我认为在日本没有哪一个手写汉字识别系统能够比得上你们的系统。”刘迎健常常对周围的同事们说:“戴老师是我的恩师,是他给了我机会。”戴汝为学术作风民主,在学生中已传为美谈。他总是尽量发挥学生的主动性、积极性和创造性,充分发挥学生自身的才能。有个学

生当年选择了国内尚未开展研究的神经元网络的课题,那时学术界有很多不同看法,认为没有前途,学生感到压力很大。这时戴汝为却坚决支持他搞下去,结果非常成功,他的研究成果被广泛引用。在解答学生的问题及进行讨论时,戴汝为总是耐心细致,一丝不苟。1991年12月,夜里两点多钟了,有个学生忽然打来电话向他请教问题,已经睡下的戴老师马上起来给他解答。半个多小时过去了,他依然在不厌其烦地认真解释着一个研究难点,使学生非常感动。如今,戴汝为的不少学生已经成为学术带头人,有的在国外成为科研工作的中坚力量。戴汝为本人于1990年获中国科学院优秀研究生导师称号。

由于在模式识别、机器智能等信息科学前沿领域作出了系统的、创造性的重要贡献,1991年他被评选为中国科学院技术科学部学部委员。他还被国家科委聘为“863”高技术信息领域智能计算机专家组成员,并任“国家智能计算机研究开发中心”首任学术委员会主任。多年来他一直担任《自动化学报》副主编、《模式识别与人工智能》杂志主编、欧洲国际性杂志《Signal Processing》的海外编委。1990年,戴汝为应邀担任在美国举行的国际“Structured Document Image Analysis”会议程序委员会的成员,1991年担任在法国举行的“Frontiers on Handwriting Recognition”会议程序委员会成员,1993年将在日本举行的 Asian Conference on Computer Vision”亦邀请他担任程序委员会成员。

几十年来,戴汝为走过了艰辛的道路。如今虽已硕果累累,但他仍持之以恒地在科研第一线辛勤耕耘。