

编者按:气象人工智能课题是中国气象局和国家自然科学基金资助项目,有高等院校、科研机构和业务部门的同志参加。它涉及到气象工作的诸多领域,其中的一些思路与设计思想具有一定开创性,为拓宽计算机的应用领域,缩短计算机应用与气象知识的距离进行了有益的探索与试验。为此,本期作专题介绍,希望它能对台站的业务工作有所推进。还需指出,介绍的这些内容只是一些初步结果,不少工作离实际业务应用还有相当差距。我们期望在不久的将来能看到水平更高、实用性更强的工作成果。

人工智能、模式识别在气象 领域应用的现状与展望

王耀生

(北京气象学院, 100081)

提 要

简要回顾了10年来人工智能、模式识别在我国气象领域进展的情况,它包括专家系统、人工神经网络、智能数据库、天气过程的计算机理解和仿真,以及卫星云图的分析识别等方面,并对今后的发展提出了一些初步设想。

关键词: 人工智能 模式识别 气象应用

引 言

气象人工智能是以人工智能技术为手段,用以推进气象科学发展的一门边缘交叉学科。由初级的气象预报专家系统形式,经过多年的发展,到现在已进入一个新的智能系统阶段。展望未来,随着气象知识的逐步积累,知识表示的不断丰富,计算机知识获取能力的日益加强,特别是随着计算机智能硬件环境的更趋完善,气象人工智能在加速气象业务现代化和提高灾害性天气预报准确率的进程中,必将发挥愈益明显的作用。

建国40年来,我国的气象事业取得了巨大的进展,特别是近10多年,在气象业务自动化及计算机的开发利用方面有了长足的进

步。

随着气象信息数量与种类的迅速增长,如何储存、管理和检索利用信息,就成为气象业务中面临的第一个问题。如何从众多的信息中将有用的部分提取出来,并按照业务要求进行组织是需要解决的又一个问题。如何有效地综合运用包括数值预报、天气图、卫星、雷达等多方面的信息、知识,提高计算机对气象信息高度综合处理能力,进一步缩小计算机处理与专家知识二者之间的距离,拓宽计算机在气象中的应用领域,加速天气预报和气象业务现代化,是当前迫切需要解决的第三个问题。

“计算机科学狭义地讲是研究符号的表示和处理,广义地讲是研究知识的表示和处

理。它超出并将继续超出纯粹的计算,帮助人们理解抽象的概念、知识的表示、问题的求解、推理和语言等,在广泛的智能活动中计算机科学起着越来越重要的作用。所有的学科都能从计算机科学中得到帮助,使用它来了解世界,解决问题,预见未来”^[1]。

人工智能和模式识别技术近年来取得了令人瞩目的成就,新理论、新方法不断涌现,它们正在改变着科学研究的方式。无论在对图象、图形的分析、识别、理解,还是在建立抽象概念、逻辑推断及问题求解等方面,提出了一系列全新的概念和方法,已经和正在解决许多过去人们期望解决的难题,从语音、文字识别、机器翻译到智能计算机的设计。将这些领域的先进思想和技术方法引入气象工作,无疑地对上述三个问题的解决将会是大有裨益的。

人工智能技术是形象知识(如预报经验)与综合推理知识(如不同预报条件的组合,形成不同的推理过程和结论)的结合,并通过不断自学习、自我完善,得出推断结论(如预报结果),而不一定需要了解某一系统(如大气)内部的微观(如物理过程的变化过程)演变。它的出现,使得利用计算机进行天气学预报方法的深化、发展成为可能,它的出现,形成了气象人工智能的交叉学科。

1 人工智能在气象领域应用的现状

人工智能技术引入气象领域在我国已经经历了10几个春秋,它大致可以区分为两个阶段。1983—1987年是第一阶段,这个阶段的主要特点是初级专家系统的普及应用。在这期间,有90%以上的省级气象台,近50%的地、市级气象台进行了气象专家系统的开发应用,它涉及到多种气象灾害的预测,如:暴雨、大风、霜冻、冰雹等。这期间在取得一定进展的同时,也暴露了它的内在局限性。自

1987年开始开发气象智能系统,经过多年发展,已逐步形成体系。

下面分别就人工智能与模式识别两方面作一回顾。

1.1 人工智能技术的应用

1.1.1 专家系统技术在我国天气预报中的应用

我国气象部门在80年代前期引进专家系统(ES)技术,1983年中国科学院大气物理研究所与吉林大学合作,开始研制“北京地区暴雨预报专家系统”。1984年王耀生等¹⁾设计的“长江中下游地区暴雨预报专家系统”,以其实用性、灵活性和多功能的特点受到广泛欢迎,目前已推广应用到中央气象台及17个省、市、自治区气象台(研究所),经历了6—7年预报业务的检验,取得了明显的效果。

该系统的又一特点是以和谐性和完备性作为机器学习的两个基本原则,在专家指导下,通过知识变换、运用归纳、类比算法,自动生成推理原则,使系统具备了一定的自学习功能。这期间陕西、湖北等省也分别独立研制了气象预报专家系统。

知识表示和知识获取作为专家系统的两个主要问题,在天气预报专家系统中更为突出,它表现在三个方面:①知识获取主要靠人工移植;②不具备记忆联想,只有微弱的自学习功能;③层次性知识不好处理。因此同一个系统外壳在不同省的应用结果差异很大。如何突破原有的专家系统框架,设计一个新的系统已成为发展的必然,1987年提出了智能预报系统的结构框架^[2,3]。这个系统的主要特点是将模式识别技术、传统人工智能与人工神经网络结合在一起。

1) 气象预报专家系统课题组,气象预报专家系统技术报告,1987年12月。

1.1.2 专家神经网络系统在暴雨预报中的应用试验^[4]

近3年来对标准误差反传(BP)网络的性能进行了试验,采用了多层网络(MLN)逐层训练算法,对隐单元层适当给予指导,并根据运行情况提供灵活的权重分配,同时又保持BP算法的优点。把多层网络的学习训练问题变成多级单层网络的学习训练。实践证明,这可以大幅度提高网络的学习训练效率。

人工神经网络(ANN)具有的自适应性和容错功能,正好弥补了ES的不足,因此把ANN与ES结合起来,可以取长补短。正是在这种思想指导下,通过对网络的相关性分析,综合统计和ES的思路,采用样本分类学习和网络多级投票。用广西、四川的10年(1981—1990)暴雨专家系统的预报因子资料进行试验,前6年用于学习,后4年作检验。结果使暴雨预报成功界限指数(CSI)=正确次数/(正确次数+漏报次数+错报次数)由原来BP网的0.50左右提高到0.60以上(附表),这表明ANN与ES相结合的专家神经网络(EANN)是用于灾害性天气预报,做好推断集成的有效工具。

附表 专家神经网络系统用四川资料检验的结果

年	1987	1988	1989	1990
错报/次	9	10	21	16
漏报/次	2	5	3	0
CSI	76%	68.7%	56.4%	65.2%

在进行暴雨预报试验中发现了一个十分有意思的现象,对全体暴雨样本集进行学习训练时,成功界限指数为0.50,其最佳训练上限只是0.60左右,约有1/4样本不符合训练要求,如果对这1/4样本与3/4样本分别进行学习训练,则成功界限指数可达0.75。这种现象在对无暴雨样本集的学习训练中也存在。这表明样本集可区分为两大类(正常类

与特殊类),它们各自有着相应的发生规律。

目前正在将人工智能技术用于暴雨强度、落区及台风的路径预报问题。

1.1.3 气象智能数据库系统^[5]

现有资料库存放的气象数据是探测数据和格点数据,不能用计算机从库中直接查询,检索有关槽、锋面、高低压中心、台风云系等有关的气象信息,尽管这些信息对气象工作者来说是十分重要的,它们与天气的变化密切相关。反映这些特征的数据是与一定的气象概念相联系的,是一种“知识型”的数据。设计的智能数据库是将数据结构与知识表示表述成统一的框架格式,从而缩小了计算机处理与气象知识二者的距离。

气象智能数据库与一般数据库相比有两个显著特点,①具有动态检索(如某类天气系统的追踪检索)和自动建模等智能检索功能;②明显的压缩信息量,一张天气图在计算机中按图形方式存放,大致为15—20kB,而在智能数据库中仅为0.5—1kB,目前正在将30年(1961—1990年)的3层(850、700、500hPa)历史天气图存入微机,预计总存储量不到40MB。

值得提出的是,正是由于上述图形数据的显著压缩,大大提高了检索速度,从30年(1961—1990年)资料中检索任意一个月、任一天的某层天气图(相当于从1000多张图中检索一张),从打入检索命令到图形屏幕显示,在386机上只需2、3秒钟,这就为气象业务应用提供了可能。

1.1.4 台风过程的计算机理解^[6,7]

一次灾害性天气的发生,往往是多种条件和因素综合影响的结果,如何很好地综合分析和应用各种气象信息,是做好灾害性天气预测的基本前提。在气象领域这种综合目前基本上还是由人来作的。

设计一种用于描述灾害性天气过程时空

结构变化的气象专用语言,提高计算机对气象信息综合分析处理能力,实现计算机对灾害性天气的自动识别、理解,进而实现基于逻辑的推断(预报)是这部分工作的主要内容。

目前的工作是以台风过程作为突破口进行试验,它的工作包括以下内容:

①建立台风过程模型;

②构造一种适合描述台风过程(模型)的类自然语言体系;

③建立与台风过程相关的气象词典及背景知识库;

④开发多用语分析器;

⑤利用气象智能数据库数据,结合台风模型,进行台风过程的计算机理解。

1.1.5 西太平洋台风路径的智能仿真(模拟)

气象仿真已有了多年的历史,早在50年代就开始了大气转盘试验,近年来计算机数据仿真已成为气象领域的一种主要方法,已经取得了不少有价值的成果,如气候成因仿真和厄尔尼诺现象的仿真等。然而在灾害性天气的仿真中出现了两方面的问题,有些由于情况太复杂(如暴雨),难于建立合适的仿真数学模型,有些又由于数学模型太复杂,一系列的参数很难设定(如台风),因而实施有困难,进展不明显。从实际应用的角度,人们对一些过程发展的细节不一定必须了解,认识和把握某些关键就可以了。就以台风路径而言,如果能够了解在什么情况台风路径将出现异常、何时将突然转向这样两个关键问题,台风路径的其它问题就显得不那么重要了。以目标系统模型为基础的计算机定性仿真正好在这种情况下可充分发挥其作用。

西太平洋台风路径的定性仿真需要经下列步骤:

①建立西太平洋台风活动的模型库(包括台风路径、强度变化、环流形势演变及要素

特征等);

②设计描述台风过程的专用计算机语言及相应的气象词典和背景知识库;

(以上两点是与计算机理解部分共用的)

③根据模型进行脚本与分场景设计,提出仿真的结构框架;

④仿真程序编制、调试;

⑤按模型要求,利用气象智能数据库,进行仿真试验;

⑥根据个例仿真结果,对模型仿真进行评估,修改模型。

这项工作经过近一年多的探索、试验,可以说前景是良好的。

1.2 模式识别应用——卫星数值云图的分析、识别试验^[8]

卫星云图在灾害性天气预报中的重要性已成为气象工作者的共识。如何从云图中提取有用的信息,实现计算机的加工处理,是本系统首先要解决的问题。我们采用数学形态学与句法模式识别相结合,将云系结构与天气系统联系起来,实现了台风过程不同阶段云系发展演变特征的计算机分析识别。

螺旋云带、密蔽云墙和台风眼是构成台风云系的3个基本特征。首先设计一套由“开”、“闭”运算组成的非线性滤波算法来滤掉卫星云图上的杂波,随后选取不同的“结构元素”作适应性分割和自动阈值判别,把台风云系从云图中抽取出来,通过“中的”运算确定台风的边界,再用“细化”算法定出台风的骨架。对边界和骨架求链码和形状数,将其按台风云系构成组成句子,储存在计算机中。结果清楚地表明在台风发展的不同阶段,上述3个基本特征的结构、形状和组成是不同的。在此基础上,还进行了台风的计算机定位试验,为客观、准确地台风定位作了一次有意义和应用前景的尝试。

2 需求分析与发展设想

气象领域人工智能、模式识别技术的应用虽然已有一定的时日,在某些方面也取得了一些进展,但总的来说这应用还是很初步的,大有开发前景。下面就气象领域的需求作一点分析,并对其发展前景提出某些设想。

2.1 卫星遥感探测问题

目前的气象卫星遥感是用扫描仪采用行扫描的办法进行,这种方法速度慢,效率低,而且时间不同步,再加上星体的态势及轨道等问题,容易造成图象的模糊与丢线等问题。如果能按照神经网络原理,引入陈列感应器件,改行扫描为帧扫描,这必将大大提高遥感探测的速度和质量。

2.2 气象数据的规范化压缩储存

气象业务是一项数据量非常惊人(每天有几百 GB),数据种类繁多,又实时性很强的工作,为了保存、管理这些资料,投入了大量的人力、财力,但是随着数据量的急速增长,利用率却越来越低,而且各种数据格式又不一致,严重地制约了它的共享与综合利用。能否实现气象数据的规范化压缩储存,通过软硬件结合,达到近两个数量级的压缩比,这不但是一个节省大笔储存开支的问题,也为广大地方台站更好地利用气象数据创造了条件。

2.3 气象数据有用信息的提取

气象图象(形)的特征提取和识别的目的是希望从浩瀚的图象(形)数据中提取出有用的信息,进而达到认识、理解的目的。由于气象图象的复杂性,单纯的统计处理、识别已遇到了原理上的困难,因此如何将统计技术与智能方法结合,就成为一个可能发展的方向,目前试验的将数学形态学与句法模式识别结合的云图识别工作就是这方面的一个探索。如何在识别基础上,对云及其发展变化做出

符合气象要求的理解,更是一项值得深入研究的课题。

2.4 气象智能数据库建设

所谓智能数据库包含两个方面的含义,数据与知识的统一以及智能检索。智能检索也就是基于知识的数据检索,它包括相似检索与演绎检索两部分,目前的工作只是完成了知识与数据结构的统一,因而大大压缩了数据量,加快了检索速度,同时也开展了静态的特征相似检索的工作。今后在实现基于知识的检索方面,还有大量工作需要做:①要建立基于灾害性天气过程模型的动态相似检索;②要建立基于演绎规则的天气过程分析、分类。在上述两项工作的基础上逐步实现计算机自动建模的工作。

2.5 气象灾害预测

实践证明气象灾害是一种最重要的自然灾害,它构成对人类的严重威胁,如何预测气象灾害,做到避害趋利,是人们长期努力的目标。

早期的气象预报专家系统,由于它知识表示的局限性及推理结构的单一,已不能满足要求。在新一代专家系统工具的基础上,开发适合气象预测要求的专家系统,并将它与人工神经网络结合,充分发挥形象处理与综合定性推理的优势,再将这与数值天气预报及其产品密切结合,是一个主要的发展方向。特别是 90 年代随着气象事业的现代化发展,信息量激增,大量的图表、数据的综合利用,需要快速作出预报,因此有必要形成一个气象灾害人工智能预测系统。

2.6 专业应用服务

气象工作本质上是一种服务性工作,由于各行各业情况各异,对气象的要求各不相同,而且这种要求也会有变化,因此要做好服务,帮助做好防灾趋利对策,就需建立服务系统。

这系统应包括相应的服务模型库、知识库(行业和气象的有关知识)、气象用语和服务用语词典,以及图形图象库等。从系统的结构可以知道,这就是要建立一个气象智能决策支持系统。

2.7 气象知识自动获取

这是一个最有兴趣,但也是最困难的问题。气象工作中有大量的各种图表,这些图表在时空上存在着内在的联系,人们通过对这些图表的考察、分析来了解大气的演变,通过对个例分析的积累,形成了概念,建立了模型,进而上升为推理,这是获取气象知识的一种途径。

以流体力学为基础,建立大气动力学、热力学方程组,通过对方程组的处理、解析以及数值分析和模拟,进而认识大气的演变规律,这是获取气象知识的又一途径。

第一类知识获取是智能型的,第二类知识获取是数理型的,这里谈的是智能型的知识获取问题。

对原始图表等主要拟采用人工神经网络通过对特征识别的记忆、联想来进行学习,个例则主要打算采用类比、归纳的学习方法。如

何将这两种学习结合,实现由此及彼,我们的设想是通过智能(定性)模拟来完成。

从长远发展来看,应该将两类知识获取途径互相结合起来。随着对大气动力学、热力学过程的深入了解,可以帮助人们建立更加合理的概念模型,而概念模型的建立和完善,又将进一步改进数值模式。

参考文献

- 1 美国国家基金会计算机研究部. ①三十年来计算机科学的发展和贡献. 模式识别与人工智能, 1990, 3(4).
- 2 王耀生. 智能预测系统设计及应用. 北京: 气象出版社, 1990年2月.
- 3 Wang yaosheng. "Intelligent Prediction System" Proceedings of the 2nd International Conference on East Asia and Western Pacific Meteorology and Climate. P561—574.
- 4 张承福. 人工神经网络在气象预报中的应用研究. 气象, 1994, 20(6).
- 5 俞善贤、王耀生等. 天气图图形数据库系统的研究. 气象, 1994, 20(6).
- 6 王耀生. 天气过程语言. 气象, 1990, 16(5).
- 7 林孔元等. 气象智能预报系统集成化问题的研究. 气象, 1994, 20(6).
- 8 王耀生等. 卫星云图的计算机分析识别. 气象, 1994, 20(6).

Application of Artificial Intelligence and Pattern Recognition in Meteorology: Development and Prospect

Wang Yaosheng

(Beijing Meteorological College 100081)

Abstract

Artificial intelligence and pattern recognition which have been used in meteorology for last 10 years are introduced concisely. It includes expert system, artificial neural network, intelligent database, computer understanding and imitation of weather process and recognition of satellite picture etc. Some advanced tentative ideas have been put forward.

Key Words: artificial intelligence pattern recognition meteorological application