

形象(直感)思维与人机结合的模式识别

戴汝为

(中国科学院自动化研究所 北京 100080)

摘 要 本文首先从启发性的论据等几个方面加以引伸,对形象(直感)思维进行探讨,并从传统的中医诊断来说明形象(直感)思维与人识别模式的思维方式,然后以人的模式识别为借鉴,讨论机器的模式识别,以综合集成(metasynthesis)为着眼点,扼要地介绍了作者所提出的“基于人工神经网络的语义句法模式识别方法”。以手写汉字识别中的子结构检测为例,阐述了有教师的学习(supervised learning)及人机结合的模式识别的重要意义。

关键词 形象思维,中国传统的医学诊断,模式识别

1 引言

80年代初,我国著名科学家钱学森同志提出创建思维科学技术部门的主张,并认为思维科学研究的突破口在于形象思维^[1]。1986年国家高技术计划把“智能计算机系统”作为信息领域的主题之一。由于对智能计算机的研讨,越来越多的人对研究形象(直感)思维的重要性有了更进一步的认识。

谈到形象思维往往使人联想到科学中的启发性论据(heuristics),就人工智能领域来说,专家系统取得了很大的进展,而启发式知识的应用及启发式搜索是专家系统不同于一般计算程序的关键之一。启发式知识是一些难以精确描述的知识,它以专家的经验为基础,以及一些直观感觉,是一种感性认识,还不是理性认识。所以我们可以对启发式知识作这样的理解:有关目前问题局况与合适的解之间的经验认识,这种经验知识难以用语言讲清楚,但又非常重要。

下面再谈谈关于如何从整体上来看待一个系统的整体性能,把握全系统的形象,而不是系统中的一枝一节。格式塔(Gestalt)心理学的观点是值得借鉴的,格式塔的一个基本特征,凡是格式塔,虽然它是由各种成份与要素构成的,但该格式塔决不等于构成它的所有成份之和。一个格式塔是一个超出于这些成份的全新整体,它是从原有的构成成份中“实现”出来的,因而它的特征与性质都是在原来的构成成份中找不到,概而言之,部份不能决定整体;“整体”的性质反过来却对“部份”的性质有着极重要的影响。必须重视整体的形象。

我们还可以从中国古代治学思想中得到启示。例如,中医里的“证”即从人体的整体状态考虑进行综合诊治,而不是头疼医头,脚疼医脚,也是讲从整体上考虑问题。另外人从眼、耳、鼻、舌等感官,获得信息,进行模式识别,把感官感觉到的信息与知识与大脑中存贮的医理模式从整体上加以比较,进行搜索与匹配找出并识别“形象”。而这种搜索与匹配的效率与情感密切相关。从以上这些方面加以引伸来探讨形象(直感)思维可以说抓住了问题的实质。

2 比喻的方法

在讨论到与形象(直感)思维有关的问题时,认为一些经验的体会是“只可意会,不可言

收稿日期:1993-11-29,这项工作得到国家科委攀登计划的资助与支持。

传”。有人就问,既然不可言传,那么怎么能理解,又如何加以研究呢?这是中国传统文化着重于“用心体会”的认知方式,存在着不容易进行交流的弱点,但绝不是无法研究。中国历史上佛教到了六祖慧能继承衣钵创立了南宗顿教,规定的教义之一是“不立文字”。讲求的是创造的直观,亦即在感受中领悟某种宇宙的规律。慧能本人是不识字但能“悟”的典范,而佛教也一直流传下来。《易传·系辞》上曾有一段很深刻的文字“书不尽言,言不尽意;……圣人立象以尽意”。前人对只可意会的东西的论述,采用的是比喻的方法。比较典型的一个例子,是范缜对形神关系的比喻,就非常精采,在中学教科书中可找到。形神的比喻是:“神之于质犹利之于刃,形之于用,犹刃之于利,利之名非刃也,刃之名非利也;然而舍利无刃,未闻刃没而利存,岂容形亡而神在?”(《神灭论》)。

前人还采用“比类取象”,“援物比类”等方法。这里我们用中医诊断疾病来对形象(直感)思维作比喻,也许可以对形象(直感)思维有进一步了解与较好的说明。从中医诊断来加以分析,首先是立“象”^[2],大夫通过多种媒体(望、闻、问、切)等感知及自身的体会,建立子模式,再形成与各种病症对应的模式类,立“象”的过程是大夫与大量的病人接触,学习与诊断的过程,病人实际上起到为大夫提供样本的作用。立“象”是十分复杂的认知过程,是大夫的实践与经验积累的过程。可以说是多维(望、闻、问、切)的自下而上的综合集成(metasynthesis)过程,这是由大夫的大脑这一系统来完成的。最终把实践经验沉积于大夫的脑中,形成表征与各种病症相对应的各种模式类的“意”。达到立象表意。如图1所示。

“立象”很重要,在此过程中,采用以象说象的办法建立一种描绘整体形象的比较抽象的象称为“意”。中医大夫对病人疾病的诊断是靠对人体的整体了解以“意”之间的相似性来加以判断。以象说象的办法可以用一个二元式来表示一个模式 P 的方式加以表达, P 包括象(用 I 表示)与意(用 E 表示)两个成份: $P = P(I, E)$ 。

其中 I : 称为模式 P 的象,包括感性成份与理性成份。

E : 称为模式 P 的意,是一种比 I 较为抽象的象也包括感性成份与理性成份。

E 的感性成份是相应的象 I 的感性成份的凝炼和浓缩。

E 的理性成份是相应的象 I 的理性成份的涵益和总结。

由于中国传统思维中认为感性与理性是相通的,所以 E 可认为是从 I 通过综合集成,即从定性到定量的综合集成所得的结果。

如图2所示,从象到意可以是一个层次,也可以用多层次形象的类比与寓意来完成。

总之,人进行模式识别的过程可以说是形象(直感)思维的主要内容之一,而综合集成在识别的过程中起着核心的作用。关于模式识别,德国的涅曼教授有如下说法:对于简单的模式,识别指的是分类;对于复杂的模式,识别指的是描述。可以说包括分类与描述两重含义。归结起

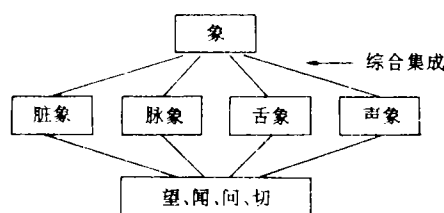


图1 中医的立象示意图

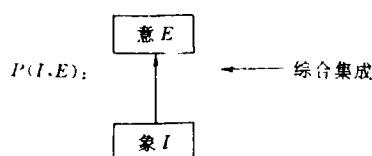


图2 从象到意的示意图

来人的模式识别有两个要点,即认知(cognition)与识别(recognition),英文的 recognition 这个字表达得较为清楚.说得更详细一点,人的模式识别是:

(1)以实践与经验为基础,通过多种媒体的认知(cognition)经反复学习,包括有教师的学习,在人脑中建立模式类,即“建模”或立“象”,并通过综合集成而从总的方面掌握“意”存贮于脑中.

(2)对新出现的模式,用一种衡量“意”之间的相似性度量进行搜索识别(recognition).判定与存贮在脑中的哪一类模式相匹配,这就需要对存贮在脑中的所有模式进行搜索.而在搜索过程中不是一个一个的顺序搜索,搜索与匹配是两个重要的环节,与人的情感有关,那种突如其来的发现,例如,突然领悟了“呀!原来是这么回事”的体验,是需要借助于情感的.当人们颇为激动的时候,他们从象到意的综合集成及搜索与匹配的效率就会提高,总之从中国传统思维中把形象(直感)思维概括为“象、意、情”的综合体现,应该说是很有道理的.

3 模式识别的语义句法方法

上节我们讨论了人的模式识别.用计算机进行模式识别所形成的软硬件技术已成为信息技术中一个组成部分.虽然也有一些人从认知的角度来研究模式识别,但到目前为止,用计算机进行模式识别的工作中,所面对的都是十分死板的模式.对于“建模”这一重要问题几乎完全忽略,把问题转为“对样本的特征抽取”.而特征抽取并没有一般的行之有效的方法.在利用计算机进行处理的前提条件下,上节所谈到关于一个模式包含两部分的构思得到充分发挥.一个模式也可定义为: $P=P(x,u)$.其中, x 表示 P 的结构信息,称为句法部分. u 表示 P 的属性等,称为语义部分.

并考虑一个模式 P 由一些称为基元(primitive)的基本元素构成,这些基元具有某些能刻画基元性质的属性,属性可以用数值表示,基元之间的相互联系用关系属性表示,由若干基元可以构成子模式,再由若干子模式构成更复杂的模式,以此类推最终构成所研究的模式.如果一个子模式由若干基元组成,而基元的属性已经知道,那么子模式的属性如何由基元的属性加以决定的办法称为语义规则或语义函数.于是就形成一个形式化的体系,可以采用一种称为“语义句法方法”的方法来描述这一体系.这种描述也包括两个部分,一是句法部分,用一个上下文无关文法或有限状态文法的导出式表示;二是语义部分,它包括基元、子模式的属性、基元与子模式间的关系属性,以及语义函数,如表1所示.

表1 语义句法的两个方面

句法部分	语义部分
树状描述	基元属性、关系属性、语义函数

在用计算机进行模式识别的发展过程中,最初采用的一般性的数学方法是统计法或决策理论法,以后又提出一种以形式语言理论为基础的句法方法.这里所谈到的语义句法方法^[3],早在1981年就提出来了,这一方法实质上是把统计方法与句法方法有效地综合在一起了.

4 人机结合的模式识别

在以往用计算机进行模式识别研究的过程中,有一点人们注意得不够,即识别模式有两种学习方法,称为有教师的学习(supervised learning)与无教师的学习(unsupervised learning)两种,前者在学习中有人参与,后者无人参与.其实用人机结合的办法来进行模式识别才是明智的,关键之处让人指点一下,大量繁琐的工作由计算机去完成,采用有教师的学习意味着人的参与是十分值得提倡的事.这一点我们还要从另一角度加以说明.在第2节中,我们讨论人

的模式识别时,从“象”到“意”是人脑加以综合集成而达到的,用计算机根本没法做到.局限于语义句法的描述而言,实际上在对语义的处理方面,如何计算语义函数也是个问题,换句话说如果一个子模式 $P(x, u)$ 由 k 个模式 P_1, \dots, P_k 个更小的子模式构成的属性记为 u , P_i 的属性为 $u_i, i=1, 2, \dots, k$. 说得简单一点所谓的语义函数,就是如何由 u_1, u_2, \dots, u_k 而得到 u . 如果一个模式由一个树状结构表示,那么问题就是如何从树的叶节点的属性逐层往上综合以求得与根节点相对应的属性. 语义函数是一个非线性的映射,这种非线性的映射计算起来比较困难. 根据近年来人工神经网络研究的进展,有人论证了在满足一定条件的情况下,一个多层人工神经网络可以实现一个非线性映射. 根据这个理论结果,我们就可以用人工神经网络来作为语义函数. 如图 3 所示. 而这种人工神经网络的参数即“权”值是通过有教师的学习来完成的,由教师把握住宏观的情况. 所以人在这一过程中起到积极的作用. 这样一来确定语义函数就有了一定的办法. 我们把模式有语义句法描述及采用人工神经网络的识别方法称为“基于人工神经网络的语义句法模式识别方法^[5]”. 从这一方法中还得到一点看法,对某些不太复杂的问题,可以用人工神经网络来实现综合集成. 上述确定语义函数的方法已在手写汉字识别中有所体现^[6],教师起了大作用. 另外,再加上下述子结构检测器的设计. 依靠这两点形成了人机结合的汉字识别方法. 识别手写汉字所遇到的困难是由于各人书写汉字时的方式与习惯不同,所以是千变万化的,计算机难以判断对于输入机器的一个汉字,有的笔划究竟是一个短划呢? 还是由于噪声形成的“毛刺”. 另外由于连笔等问题把笔划当成基本元素不实际. 所以把字根(偏旁部首)作为基本单元,称为子结构. 通过一些运算把子结构检测出来,与子结构对应的运算是靠人确定的. 一个子结构由若干笔划组成,检测到的子结构由畸变的笔划构成,具有模糊性,把这些模糊了的笔划的属性 u_1, u_2, \dots, u_k 送到一个类似于感知器(perceptron)的网络中,然后根据一系列的样本来进行学习,并经过一定的非线性处理,求得与子结构对应的属性(用一个向量表示)如图 4 所示,例如对于字根“木”,把左偏旁为“木”的字作为正样本,把左偏旁不是“木”的字作为负样本,把大量正、负样本输入计算机,教师的作用是把把握宏观状况. 当一个样本输入时确定网络的输出应该是“正”或“负”,用一个算法对参数加以调整,并给出一定的准则以确定在 n 个参数中,有较大影响的 P_1, P_2, \dots, P_m m 个,则这个量可以定义为子结构的属性. 该训练过的网络就起到语义函数的作用.



图 3 语义函数的构建

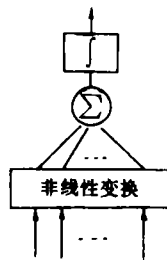


图 4 求与子结构对应的属性

以上介绍的是十分具体的识别问题,从中看出对于不太复杂的问题,我们从工程技术的角度,可以采用人工神经网络及有教师的学习方法,或人机结合方法来实现综合集成(metasynthesis). 它告诉我们形象思维的重要性,也启发我们在研究更复杂的巨系统问题时,该如何去开拓前进的道路.

致谢 在完成本文的过程中,作者多次从钱学森先生的来信中得到启示、鼓励与帮助,在此表示衷心的感谢!

(下转第 118 页)

5 结束语

本文提出一种基于推广卡尔曼滤波的前馈神经网络快速学习方法,有效地克服了 BP 算法收敛速度慢且常常收敛于局部极值的缺陷,仿真计算表明,该算法收敛快、数值稳定性好、精度高、简单实用,是前馈神经网络一种有效的学习算法.该方法把权系数和阈值作为状态变量来处理,所以,其计算量将随隐层节点数的增加而增大,为此,可采用 U-D 分解技术等措施以进一步提高计算效率,改善学习效果,这一工作将另文介绍.

参 考 文 献

- 1 Rumelhart D E, McClelland J. *Parallel Distributed Processing*. Cambridge, MA: MIT Press, 1986, 1.
- 2 Jacobs R A. Increased Rates of Convergence through Learning Rate Adaptation. *Neural Networks*, 1988, 1: 295-307.
- 3 Singhal S, Wu L. Training Feed-Forward Networks with the Extended Kalman Algorithm. *Proc ICASSP*, 1989, 1187-1190.
- 4 Ruchti T L, Brown R H. Estimation of Artificial Neural Network Parameters for Nonlinear System Identification. *Proc of 31st Conf on Decision and Control*, 1992, 2728-2733.
- 5 Anderson B D O, Moore J B. *Optimal Filtering*. Prentice Hall, 1979.
- 6 Chen S, Billings S A. Neural Networks for Nonlinear Dynamic System Modelling and Identification. *Int J Control*, 1992, 56(2): 319-346.

(张友民,男,30岁,硕士.主要研究领域为滤波,辨识与随机控制,神经网络及其在控制中的应用,故障诊断与容错控制,控制系统中的并行处理等.)

(上接第79页)

参 考 文 献

- 1 钱学森主编.关于思维科学.上海人民出版社,1986.
- 2 王前,刘庚祥.从中医取象看中国传统抽象思维. *哲学研究*, 1993, (4): 45-50.
- 3 戴汝为.从定性到定量综合集成技术. *模式识别与人工智能*, 1991, 4(1): 5-10.
- 4 DAI Ruwei, Fu Jingsun. Semantic Syntax-Directed Translation for Pictorial Pattern Recognition. *Purdue Univ Tech Rep*, 1981, TR-EE-81-38.
- 5 DAI Ruwei. A Connectionist Syntactic Semantic Approach for Pictorial Pattern Recognition. *Proc InfoScience'93*, Seoul, Korea, (Invited Paper), 1993, 233-240.
- 6 刘迎建,戴汝为,张立清.基于神经网络的手写汉字特征选择. *模式识别与人工智能*, 1992, 5(1): 37-43.

THINKING WITH IMAGERY (INTUITIVE) AND PATTERN RECOGNITION BY MAN MACHINE COOPERATION

DAI Ruwei

(*Institute of Automation, Chinese Academy of Sciences*)

Abstract The thinking with imagery (intuitive) is investigated from chinese traditional medical diagnosis. Besides, the problems of pattern recognition by machine, and a brief introduction on "connectionist syntactic semantic approach of pattern recognition" proposed by the author are also given.

Key words thinking with imagery, chinese traditional medical diagnosis, pattern recognition

(戴汝为,男,61岁,研究员,院士.所从事的研究领域为人工智能,模式识别.)